

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : A01H 1/00, C07H 21/04, C07K 14/00, C12N 5/04, 5/10, C12P 19/34, C12Q 1/68		A1	(11) International Publication Number: WO 98/30083 (43) International Publication Date: 16 July 1998 (16.07.98)
(21) International Application Number: PCT/US98/00615 (22) International Filing Date: 9 January 1998 (09.01.98)		(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Priority Data: 08/781,734 10 January 1997 (10.01.97) US		Published <i>With international search report.</i>	
(71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 22nd floor, 300 Lakeside Drive, Oakland, CA 94612 (US).			
(72) Inventors: SHEN, Kathy; 44228 Country Club Drive, El Macero, CA 95618 (US). MEYERS, Blake; 904 Drake Drive, Davis, CA 95616 (US). MICHELMORE, Richard, W.; 36757 Russel Boulevard, Davis, CA 95616 (US).			
(74) Agents: EINHORN, Gregory, P. et al.; Townsend and Townsend and Crew LLP, 8th floor, Two Embarcadero Center, San Francisco, CA 94111 (US).			
(54) Title: RG NUCLEIC ACIDS FOR CONFERRING DISEASE RESISTANCE TO PLANTS			
(57) Abstract <p>The present invention provides RG nucleic acids and proteins which confer disease resistance to plants. The nucleic acids can be used to produce transgenic plants resistant to pests. Antibodies to proteins of the invention are also provided.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakhstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

RG NUCLEIC ACIDS FOR CONFERRING DISEASE RESISTANCE TO PLANTS

5

The present application is a continuation-in-part application ("CIP") of U.S. Patent Application Serial No. ("USSN") 08/781,734, filed January 10, 1997. The 10 aforementioned application is explicitly incorporated herein by reference in its entirety and for all purposes.

This invention was made with Government support under Grant Nos. 92-37300-7547 and 95-37300-1571, awarded by the United States Department of Agriculture. 15 The Government has certain rights in this invention.

FIELD OF THE INVENTION

The present invention relates generally to plant molecular biology. In particular, it relates to nucleic acids and methods for conferring pest resistance in plants. 20 particularly lettuce.

BACKGROUND OF THE INVENTION

Recently, several resistance genes have been cloned by several groups from several plants. Many of these genes are sequence related. The derived amino acid 25 sequences of the most common class, *RPS2*, *RPM1* (bacterial resistances in *Arabidopsis* (Mindrinos *et al.* *Cell* 78:1089-1099 (1994)); Bent *et al.* *Science* 265:1856-1860 (1994); Grant *et al.*, *Science* 269:843-846 (1995)), *L6* (fungal resistance in flax; Lawrence, *et al.*, *The Plant Cell* 7:1195-1206 (1995)), and *N*, (virus resistance in tobacco; Whitham, *et al.*, *Cell* 78:1101-1115 (1994); and U.S. Patent No. 5,571,706), all contain leucine-rich 30 repeats (LRR) and nucleotide binding sites (NBS).

The NBS is a common motif in several mammalian gene families encoding signal transduction components (e.g., *Ras*) and is associated with ATP/GTP-binding sites.

The NBS is a common motif in several mammalian gene families encoding signal transduction components (e.g., *Ras*) and is associated with ATP/GTP-binding sites.

LRR domains can mediate protein-protein interactions and are found in a variety of proteins involved in signal transduction, cell adhesion and various other 5 functions. LRRs are leucine rich regions often comprising 20-30 amino acid repeats where leucine and other aliphatic residues occur periodically. LRRs can function extracellularly or intracellularly.

Since the onset of civilization, plant diseases have had catastrophic effects on crops and the well-being of the human population. Plant diseases continue to effect 10 enormous human and economic costs. An increasing human population and decreasing amounts of arable land make all approaches to preventing and treating plant pathogen destruction critical. The ability to control and enhance a plant's protective responses against pathogens would be of enormous benefit. Tissue-specific and temporal control of mechanisms responsible for plant cell death would also be of great practical and economic 15 value. The present invention fulfills these and other needs.

What is needed in the art are plant disease resistance genes and means to create transgenic disease resistance plants, particularly in lettuce. Further, what is needed in the art is a means to DNA fingerprint cultivars and germplasm with respect to their disease resistance haplotypes for use in plant breeding programs. The present invention 20 provides these and other advantages.

SUMMARY OF THE INVENTION

The present invention provides isolated nucleic acid constructs. These constructs comprise an RG (resistance gene) polynucleotide which encodes an RG 25 polypeptide having at least 60% sequence identity to an RG polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, and an RG4 polypeptide. RG1, RG2, RG3, RG4, and the like, represent individual "RG families." Each "RG family," as defined herein, is a group of polypeptide sequences that have at least 60% amino acid sequence identity. Individual members of an RG family, 30 *i.e.*, individual species of the genus, typically map to the same genomic locus. The invention provides for constructs comprising nucleotides encoding the RG families of the

invention, which can include sequences encoding a leucine rich region (LRR), and/or a nucleotide binding site (NBS), or both.

The invention provides for an isolated nucleic acid construct comprising an RG polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide from an RG family selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide. In alternative embodiments, the nucleic acid construct comprises an RG polynucleotide which encodes an RG polypeptide comprising an leucine rich region (LRR), or, an RG polypeptide comprising a nucleotide binding site (NBS). The nucleic acid construct can comprise a polynucleotide which is a full length gene. In another embodiment, the nucleic acid construct encodes a fusion protein.

In one embodiment, the nucleic acid construct comprises a sequence encoding an RG1 polypeptide. The RG1 polypeptide can be encoded by a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 and SEQ ID NO:137 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).

In another embodiment, the nucleic acid construct comprises a sequence encoding an RG2 polypeptide. The RG2 polypeptide can be encoded by a polynucleotide sequence selected from the group consisting of: SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

In other embodiments, the nucleic acid construct comprises a RG3 sequence (SEQ ID NO:68) encoding an RG3 polypeptide (SEQ ID NO:138) (RG3). In other embodiments, the nucleic acid construct comprises an RG4 sequence (SEQ ID NO:69) encoding an RG4 polypeptide (SEQ ID NO:139) (RG4).

5 In other embodiments, the nucleic acid construct comprises a RG5 sequence (SEQ ID NO:134) encoding an RG5 polypeptide (SEQ ID NO:135). The RG5 polypeptide can be encoded by a polynucleotide sequence as set forth in SEQ ID NO:134.

10 The invention also provides for a nucleic acid construct which comprises an RG7 sequence encoding an RG7 polypeptide. The RG7 polypeptide can be encoded by a polynucleotide sequence as set forth in SEQ ID NO:136.

15 In further embodiments, the nucleic acid construct can further comprise a promoter operably linked to the RG polynucleotide. In alternative embodiments, the promoter can be a plant promoter; a disease resistance promoter; a lettuce promoter; a constitutive promoter; an inducible promoter; or, a tissue-specific promoter. The nucleic acid construct can comprise a promoter sequence from an RG gene linked to a heterologous polynucleotide.

20 The invention also provides for a transgenic plant comprising a recombinant expression cassette comprising a promoter operably linked to an RG polynucleotide. The expression cassette can comprise a plant promoter or a viral promoter; the plant promoter can be a heterologous promoter. In one embodiment, the transgenic plant is lettuce. In alternative embodiments, the transgenic plant comprises an expression cassette which includes an RG polynucleotide selected from the group consisting of SEQ ID NO:1 (RG1A); SEQ ID NO:2 and SEQ ID NO:137 (RG1B); SEQ ID NO: 3 (RG1C); SEQ ID NO:4 (RG1D); SEQ ID NO:5 (RG1E); SEQ ID NO:6 (RG1F); SEQ ID NO:7 (RG1G);
25 SEQ ID NO:8 (RG1H); SEQ ID NO:9 (RG1I) and SEQ ID NO:10 (RG1J); SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104

(RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W); SEQ ID NO:68 (RG3); SEQ ID NO:69 (RG4); SEQ ID NO:134 (RG5); or SEQ ID NO:136 (RG7).

The invention provide for a transgenic plant comprising an expression cassette comprising an RG polynucleotide which can encode an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO:13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), or SEQ ID NO:20 (RG1J); or, an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W); an RG4 polypeptide as set forth by SEQ ID NO:72; an RG5 polypeptide with a sequence as set forth by SEQ ID NO:135; or, an RG7 polypeptide.

The invention also provides for a method of enhancing disease resistance in a plant, the method comprising introducing into the plant a recombinant expression cassette comprising a promoter functional in the plant and operably linked to an RG polynucleotide sequence. In this method, the plant can be a lettuce plant; and, the RG polynucleotide can encode an RG polypeptide selected from the group consisting of an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO:13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID

NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), or SEQ ID NO:20 (RG1J); or, an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:72; SEQ ID NO:74; SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W). In this method, the promoter can be a plant disease resistance promoter, a tissue-specific promoter, a constitutive promoter, or an inducible promoter.

The invention also provides for a method of detecting RG resistance genes in a nucleic acid sample, the method comprising: contacting the nucleic acid sample with an RG polynucleotide to form a hybridization complex; and, wherein the formation of the hybridization complex is used to detect the RG resistance gene in the nucleic acid sample. In this method, the RG polynucleotide can be an RG1 polynucleotide, an RG2 polynucleotide, an RG3 polynucleotide, an RG4 polynucleotide, an RG5 polynucleotide or an RG7 polynucleotide. In this method, the RG resistance gene can be amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide, and, the RG resistance gene can be amplified by the polymerase chain reaction. In one embodiment, the RG polynucleotide is labeled.

The invention further provides for an RG polypeptide having at least 60% sequence identity to a polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

A further understanding of the nature and advantages of the present invention may be realized by reference to the remaining portions of the specification, the figures and claims.

All publications, patents and patent applications cited herein are hereby
5 expressly incorporated by reference for all purposes.

DETAILED DESCRIPTION OF THE INVENTION

This invention relates to families of RG genes, particularly from *Lactuca sativa*. Nucleic acid sequences of the present invention can be used to confer resistance in
10 plants to a variety of pests including viruses, fungi, nematodes, insects, and bacteria. Sequences from within the RG genes can be used to fingerprint cultivars or germplasm for
the presence of desired resistance genes. Promoters of RG genes can be used to drive
heterologous gene expression under conditions in which RG genes are expressed. Further,
15 the present invention provides RG proteins and antibodies specifically reactive to RG
proteins. Antibodies to RG proteins can be used to detect the type and amount of RG
protein expressed in a plant sample.

The present invention has use over a broad range of types of plants, including species from the genera *Cucurbita*, *Rosa*, *Vitis*, *Juglans*, *Fragaria*, *Lotus*,
Medicago, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*,
20 *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*,
Hyoscyamus, *Lycopersicon*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*,
Ciahorium, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Heterocallis*, *Nemesis*,
Pelargonium, *Panieum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*,
Browaalia, *Glycine*, *Pisum*, *Phaseolus*, *Lolium*, *Oryza*, *Zea*, *Avena*, *Hordeum*, *Secale*,
25 *Triticum*, and, *Sorghum*. In particularly preferred embodiments, species from the family
Compositae and in particular the genus *Lactuca* are employed such as *L. sativa* and such
subspecies as *crispa*, *longifolia*, and *asparagina*.

The nucleic acids of the present invention can be used in marker-aided selection. Marker-aided selection does not require the complete sequence of the gene or
30 precise knowledge of which sequence confers which specificity. Instead, partial sequences can be used as hybridization probes or as the basis for oligonucleotide primers to amplify nucleic acid, e.g., by PCR. Partial sequences can be used in other methods, such as to

follow the segregation of chromosome segments containing resistance genes in plants. Because the RG marker is the gene itself, there can be negligible recombination between the marker and the resistance phenotype. Thus, RG polynucleotides of the present invention provide an optimal means to DNA fingerprint cultivars and wild germplasm with respect to their disease resistance haplotypes. This can be used to indicate which germplasm accessions and cultivars carry the same resistance genes. At present, selection of plants (e.g., lettuce) for resistance to some diseases is slow and difficult. But linked markers allow indirect selection for such resistance genes. Moreover, RG markers also allow resistance genes to be identified and combined in a manner that would not otherwise be possible. Numerous accessions have been identified that provide resistance to all isolates of downy mildew (*Bremia lactucae*). However, without molecular markers it is impossible to combine such resistances from different sources. The nucleic acid sequences of the invention provide for a fast and convenient means to identify and combine resistances from different sources. The RG markers of the invention can also be used to identify recombinants that have new combinations of resistance genes in *cis* on the same chromosome.

In addition, RG markers may allow the identification of the Mendelian factors determining traits, such as field resistance to downy mildew. Once such markers have been identified, they will greatly increase the ease with which field resistance can be transferred between lines and combined with other resistances.

In another application, primers to RG sequences can be also designed to amplify sequences that are conserved in multiple RG family members. This gives genetic information on multiple RG family members. Alternatively, one or more primers can be made to sequences unique to a single resistance gene genus or a single RG specie. This allows an analysis of individual family groups (an RG genus) or an individual family member (a specie). Primers made to individual RGs at the edge of each cluster can be used to select for recombinants within the cluster. This minimizes the amount of linkage drag during introgression. Classical and molecular genetics has shown that pest resistance genes tend to be clustered in the genome. Pest resistance loci comprise arrays of genes and exhibit a variety of complex haplotypes rather than being simple alternate allelic forms. Pest resistance is conferred by families, or genera, of related RG sequences, individual members, or species, of which have evolved to have a different specificity.

Oligonucleotide primers can be designed that amplify members from multiple haplotypes, or genuses, or amplify only members of one genus, or only amplify an individual specie. This will provide codominant information and allow heterozygotes to be distinguished from homozygotes.

5 Further, comparison of RG sequences will allow a determination of which sequences are critical for resistance and will ultimately lead to engineering resistance genes with new specificities. Resistance gene sequences were not previously available for lettuce. Marker-aided selection will greatly increase the precision and speed of breeding for disease resistance. Transgenic approaches will allow pyramiding of resistance genes
10 into a single Mendelian unit, transfer between sexually-incompatible species, substitute for conventional backcrossing procedures, and allow expression of other genes in parallel with resistance genes.

The RG polynucleotides also have utility in the construction of disease resistant transgenic plants. This avoids lengthy and sometimes difficult backcrossing
15 programs currently necessary for introgression of resistance. It is also possible to transfer resistance polynucleotides between sexually-incompatible species, thereby greatly increasing the germplasm pool that can be used as a source of resistance genes. Cloning of multiple RG sequences in a single cassette will allow pyramiding of genes for resistance against multiple isolates of a single pathogen such as downy mildew or against multiple
20 pathogens. Once introduced, such a cassette can be manipulated by classical breeding methods as a single Mendelian unit.

Transgenic plants of the present invention can also be constructed using an RG promoter. The promoter sequences from RG sequences of the invention can be used with RG genes or heterologous genes. Thus, RG promoters can be used to express a
25 variety of genes in the same temporal and spatial patterns and at similar levels to resistance genes.

Nucleic acids of the Invention and Their Preparation

RG Polynucleotide Families

30 The present invention provides isolated nucleic acid constructs which comprise an RG polynucleotide. In alternative embodiments, the RG polynucleotide is at least 18 nucleotides in length, typically at least 20, 25, or 30 nucleotides in length, more

typically at least 100 nucleotides in length, generally at least 200 nucleotides in length, preferably at least 300 nucleotides in length, more preferably at least 400 nucleotides in length, and most preferably at least 500 nucleotides in length.

In particularly preferred embodiments, the RG polynucleotide encodes a RG protein which confers resistance to plant pests. This RG protein can be longer, equivalent, or shorter than the RG protein encoded by an RG gene. In various embodiments, an RG polynucleotide can hybridize under stringent conditions to members of an RG family (an RG genus); *e.g.*, it can hybridize to a member of the RG1 RG family, such as an RG1 polynucleotide selected from the group consisting of: SEQ ID NO:1 (RG1A); SEQ ID NO:2 and SEQ ID NO:137 (RG1B); SEQ ID NO: 3 (RG1C); SEQ ID NO:4 (RG1D); SEQ ID NO:5 (RG1E); SEQ ID NO:6 (RG1F); SEQ ID NO:7 (RG1G); SEQ ID NO:8 (RG1H); SEQ ID NO:9 (RG1I) and SEQ ID NO:10 (RG1J).

In other embodiments, the polynucleotide can also hybridize under stringent conditions to a member of the RG2 family; such as an RG2 polynucleotide selected from the group consisting of: SEQ ID NO:21 and SEQ ID NO:27 (RG2A); SEQ ID NO:23 and SEQ ID NO:28 (RG2B); SEQ ID NO:29 (RG2C); SEQ ID NO:30 (RG2D); SEQ ID NO:31 (RG2E); SEQ ID NO:32 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:34 (RG2H); SEQ ID NO:35 (RG2I); SEQ ID NO:36 (RG2J); SEQ ID NO:37 (RG2K); SEQ ID NO:38 (RG2L); SEQ ID NO:39 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).

In alternative embodiments, each RG2 gene can also include an AC15 sequence which hybridizes under stringent conditions to a polynucleotide selected from the group consisting of: SEQ ID NO:56 (AC15-2A); SEQ ID NO:57 (AC15-2B); SEQ ID NO:58 (AC15-2C); SEQ ID NO:59 (AC15-2D); SEQ ID NO:60 (AC15-2E); SEQ ID NO:61 (AC15-2G); SEQ ID NO:62 (AC15-2H); SEQ ID NO:63 (AC15-2I); SEQ ID

NO:64 (AC15-2J); SEQ ID NO:65 (AC15-2L); SEQ ID NO:66 (AC15-2N); SEQ ID NO:67 (AC15-2O).

In other embodiments, an RG polynucleotide can hybridize under stringent conditions to an RG3 (SEQ ID NO:68), an RG4 (SEQ ID NO:69), and RG5 (SEQ ID NO:135), and an RG7 (SEQ ID NO:137), RG family member.

The present invention further provides nucleic acid constructs which comprise an RG polynucleotide which encodes RG polypeptides from various RG families; such as an RG polypeptide having at least 60% sequence identity to an RG polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, and RG4 polypeptide, and RG5 polypeptide, and an RG7 polypeptide.

Exemplary RG1 polypeptides have the sequences shown in SEQ ID NO:2 (RG1A), SEQ ID NO:4 (RG1B), SEQ ID NO:6 (RG1C), SEQ ID NO:8 (RG1D), SEQ ID NO:10 (RG1E), SEQ ID NO:12 (RG1F), SEQ ID NO:14 (RG1G), SEQ ID NO:16 (RG1H), SEQ ID NO:20 (RG1J). Exemplary RG2 polypeptides have the sequences shown in SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

An exemplary RG3 polypeptide has the sequence shown in SEQ ID NO:138. An exemplary RG4 polypeptide has the sequence shown in SEQ ID NO:139. RG polynucleotides will have at least 60% identity, more typically at least 65% identity, generally at least 70% identity, and preferably at least 75% identity, more preferably at least 80% identity, and most preferably at least 85%, 90%, or 95% identity at the deduced amino acid level. The regions where substantial identity is assessed can be inclusive or exclusive of the nucleotide binding site or the leucine rich region.

Vectors and Transcriptional Control Elements

The invention, providing methods and reagents for making novel species and genuses of RG nucleic acids described herein, further provides methods and reagents for expressing these nucleic acids using novel expression cassettes, vectors, transgenic plants and animals, using constitutive and inducible transcriptional and translational *cis*- (e.g., promoters and enhancers) and *trans*-acting control elements.

The expression of natural, recombinant or synthetic plant disease resistance polypeptide-encoding or other (*i.e.*, antisense, ribozyme) nucleic acids can be achieved by operably linking the coding region a promoter (that can be plant-specific or not, 10 constitutive or inducible), incorporating the construct into an expression cassette (such as an expression vector), and introducing the resultant construct into an *in vitro* reaction system or a suitable host cell or organism. Synthetic procedures may also be used. Typical expression systems contain, in addition to coding or antisense sequence, 15 transcription and translation terminators, polyadenylation sequences, transcription and translation initiation sequences, and promoters useful for transcribing DNA into RNA. The expression systems optionally at least one independent terminator sequence, sequences permitting replication of the cassette *in vivo*, *e.g.*, plants, eukaryotes, or prokaryotes, or a combination thereof, (*e.g.*, shuttle vectors) and selection markers for the selected expression system, *e.g.*, plant, prokaryotic or eukaryotic systems. To ensure proper 20 polypeptide expression under varying conditions, a polyadenylation region at the 3'-end of the coding region can be included (see Li (1997) *Plant Physiol.* 115:321-325, for a review of the polyadenylation of RNA in plants). The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA (*e.g.*, using *Agrobacterium tumefaciens* T-DNA replacement vectors, see *e.g.*, Thykjaer (1997) *Plant Mol Biol.* 35:523-530; using a plasmid containing a gene of interest flanked by 25 *Agrobacterium* T-DNA border repeat sequences; Hansen (1997) "T-strand integration in maize protoplasts after codelivery of a T-DNA substrate and virulence genes," *Proc. Natl. Acad. Sci. USA* 94:11726-11730.

To identify the promoters, the 5' portions of the clones described here are 30 analyzed for sequences characteristic of promoter sequences. For instance, promoter sequence elements include the TATA box consensus sequence (TATAAT), which is usually 20 to 30 base pairs upstream of the transcription start site. In plants, further

upstream from the TATA box, at positions -80 to -100, there is typically a promoter element with a series of adenines surrounding the trinucleotide G (or T) N G (see, e.g., Messing, in *Genetic Engineering in Plants*, pp. 221-227, Kosage, Meredith and Hollaender, eds. 1983). If proper polypeptide expression is desired, a polyadenylation region at the 3'-end of the RG coding region should be included. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from viral genes, such as T-DNA.

The nucleic acids of the invention can be expressed in expression cassettes, vectors or viruses which are transiently expressed in cells using, for example, episomal expression systems (e.g., cauliflower mosaic virus (CaMV) viral RNA is generated in the nucleus by transcription of an episomal minichromosome containing supercoiled DNA, Covey (1990) *Proc. Natl. Acad. Sci. USA* 87:1633-1637). Alternatively, coding sequences can be inserted into the host cell genome becoming an integral part of the host chromosomal DNA.

Selection markers can be incorporated into expression cassettes and vectors to confer a selectable phenotype on transformed cells and sequences coding for episomal maintenance and replication such that integration into the host genome is not required. For example, the marker may encode biocide resistance, such as antibiotic resistance, particularly resistance to chloramphenicol, kanamycin, G418, bleomycin, hygromycin, or herbicide resistance, such as resistance to chlorosulfuron or Basta, to permit selection of those cells transformed with the desired DNA sequences, see for example, Blondelet-Rouault (1997) *Gene* 190:315-317; Aubrecht (1997) *J. Pharmacol. Exp. Ther.* 281:992-997. Because selectable marker genes conferring resistance to substrates like neomycin or hygromycin can only be utilized in tissue culture, chemoresistance genes are also used as selectable markers *in vitro* and *in vivo*. See also, Mengiste (1997) "High-efficiency transformation of *Arabidopsis thaliana* with a selectable marker gene regulated by the T-DNA 1' promoter," *Plant J.* 12:945-948, showing that the 1' promoter is an attractive alternative to the cauliflower mosaic virus (CaMV) 35S promoter for the generation of T-DNA insertion lines, the 1' promoter may be especially beneficial for the secondary transformation of transgenic strains containing the 35S promoter to exclude homology-mediated gene silencing.

The endogenous promoters from the RG genes of the present invention can be used to direct expression of the genes. These promoters can also be used to direct expression of heterologous structural genes. The promoters can be used, for example, in recombinant expression cassettes to drive expression of genes conferring resistance to any number of pathogens or pests, including fungi, bacteria, and the like.

5 *Constitutive Promoters*

In construction of recombinant expression cassettes, vectors, transgenics, of the invention, a promoter fragment can be employed to direct expression of the desired gene in all tissues of a plant or animal. Promoters that drive expression continuously under physiological conditions are referred to as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation. Examples of constitutive promoters include those from viruses which infect plants, such as the cauliflower mosaic virus (CaMV) 35S transcription initiation region; the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*; the promoter of the tobacco mosaic virus; and, other transcription initiation regions from various plant genes known to those of skill. See also Holtorf (1995) "Comparison of different constitutive and inducible promoters for the overexpression of transgenes in *Arabidopsis thaliana*," *Plant Mol. Biol.* 29:637-646.

10 *Inducible Promoters*

Alternatively, a plant promoter may direct expression of the plant disease resistance nucleic acid of the invention under the influence of changing environmental conditions or developmental conditions. Examples of environmental conditions that may effect transcription by inducible promoters include pathogenic attack, anaerobic conditions, elevated temperature, drought, or the presence of light. Such promoters are referred to herein as "inducible" promoters. For example, the invention incorporates the drought-inducible promoter of maize (Busk (1997) *supra*); the cold, drought, and high salt inducible promoter from potato (Kirch (1997) *Plant Mol. Biol.* 33:897-909).

Embodiments of the invention also incorporate use of plant promoters which are inducible upon injury or infection to express the invention's plant disease resistance (RG) polypeptides. Various embodiments include use of, e.g., the promoter for a tobacco (*Nicotiana tabacum*) sesquiterpene cyclase gene (EAS4 promoter), which is expressed in wounded leaves, roots, and stem tissues, and upon infection with microbial pathogens (Yin

(1997) *Plant Physiol.* 115(2):437-451); the ORF13 promoter from *Agrobacterium rhizogenes* 8196, which is wound inducible in a limited area adjacent to the wound site (Hansen (1997) *Mol. Gen. Genet.* 254:337-343); the Shpx6b gene promoter, which is a plant peroxidase gene promoter induced by microbial pathogens (demonstrated using a 5 fungal pathogen, see Curtis (1997) *Mol. Plant Microbe Interact.* 10:326-338); the wound-inducible gene promoter wun1, derived from potato (Siebertz (1989) *Plant Cell* 1:961-968); the wound-inducible *Agrobacterium pmas* gene (mannopine synthesis gene) promoter (Guevara-Garcia (1993) *Plant J.* 4:495-505).

Alternatively, plant promoters which are inducible upon exposure to plant 10 hormones, such as auxins, are used to express the nucleic acids of the invention. For example, the invention can use the auxin-response elements E1 promoter fragment (AuxREs) in the soybean (*Glycine max L.*) (Liu (1997) *Plant Physiol.* 115:397-407); the auxin-responsive Arabidopsis GST6 promoter (also responsive to salicylic acid and hydrogen peroxide) (Chen (1996) *Plant J.* 10: 955-966); the auxin-inducible parC 15 promoter from tobacco (Sakai (1996) 37:906-913); a plant biotin response element (Streit (1997) *Mol. Plant Microbe Interact.* 10:933-937); and, the promoter responsive to the stress hormone abscisic acid (Sheen (1996) *Science* 274:1900-1902).

Plant promoters which are inducible upon exposure to chemicals reagents 20 which can be applied to the plant, such as herbicides or antibiotics, are also used to express the nucleic acids of the invention. For example, the maize In2-2 promoter, activated by benzenesulfonamide herbicide safeners, can be used (De Veylder (1997) *Plant Cell Physiol.* 38:568-577); application of different herbicide safeners induces distinct gene expression patterns, including expression in the root, hydathodes, and the shoot apical meristem. Coding sequence can be under the control of, e.g., a tetracycline-inducible 25 promoter, e.g., as described with transgenic tobacco plants containing the *Avena sativa L.* (oat) arginine decarboxylase gene (Masgrau (1997) *Plant J.* 11:465-473); or, a salicylic acid-responsive element (Stange (1997) *Plant J.* 11:1315-1324. Using chemically- (e.g., hormone- or pesticide-) induced promoters, harvesting of fruits and plant parts would be greatly facilitated. A chemical which can be applied to the transgenic plant in the field and 30 induce expression of a polypeptide of the invention throughout all or most of the plant would make a environmentally safe defoliant or herbicide. Thus, the invention also provides for transgenic plants containing an inducible gene encoding for the RG

polypeptides of the invention whose host range is limited to target plant species, such as weeds or crops before, during or after harvesting.

Abscission promoters are activated upon plant ripening, such as fruit ripening, and are especially useful incorporated in the expression systems (*e.g.*, expression cassettes, vectors) of the invention. In some embodiments, when a plant disease resistant polypeptide-encoding nucleic acid is under the control of such a promoter, rapid cell death, induced by expression of the invention's polypeptide, can accelerate and/or accentuate abscission, increasing the efficiency of the harvesting of fruits or other plant parts, such as cotton, and the like. Induction of rapid cell death at this time would accelerate separation of the fruit from the plant, greatly augmenting harvesting procedures. See, *e.g.*, Kalaitzis (1997) *Plant Physiol.* 113:1303-1308, discussing tomato leaf and flower abscission; Payton (1996) *Plant Mol. Biol.* 31:1227-1231, discussing ethylene receptor expression regulation during fruit ripening, flower senescence and abscission; Koehler (1996) *Plant Mol. Biol.* 31:595-606, discussing the gene promoter for a bean abscission cellulase; Kalaitzis (1995) *Plant Mol. Biol.* 28: 647-656, discussing cloning of a tomato polygalacturonase expressed in abscission; del Campillo (1996) *Plant Physiol.* 111:813-820, discussing pedicel breakstrength and cellulase gene expression during tomato flower abscission.

Tissue-Specific Promoters

Tissue specific promoters are transcriptional control elements that are only active in particular cells or tissues. Plant promoters which are active only in specific tissues or at specific times during plant development are used to express the nucleic acids of the invention. Examples of promoters under developmental control include promoters that initiate transcription only in certain tissues, such as leaves, roots, fruit, seeds, ovules, pollen, pistils, or flowers. Such promoters are referred to as "tissue specific". The operation of a promoter may also vary depending on its location in the genome. Thus, an inducible promoter may become fully or partially constitutive in certain locations.

For example, a seed-specific promoter directs expression in seed tissues. Such promoters may be, for example, ovule-specific, embryo-specific, endosperm-specific, integument-specific, seed coat-specific, or some combination thereof. A leaf-specific promoter has been identified in maize, Busk (1997) *Plant J.* 11:1285-1295. The ORF13 promoter from *Agrobacterium rhizogenes* exhibits high activity in roots (Hansen (1997) *supra*). A maize pollen-specific promoter has been identified in maize (Guerrero (1990)

Mol. Gen. Genet. 224:161-168). A tomato promoter active during fruit ripening, senescence and abscission of leaves and, to a lesser extent, of flowers can be used (Blume (1997) *Plant J.* 12:731-746). A pistol specific promoter has been identified in the potato (*Solanum tuberosum L.*) SK2 gene, encoding a pistil-specific basic endochitinase (Ficker 5 (1997) *Plant Mol. Biol.* 35:425-431). The Blec4 gene from pea (*Pisum sativum cv. Alaska*) is active in epidermal tissue of vegetative and floral shoot apices of transgenic alfalfa, making it a useful tool to target the expression of foreign genes to the epidermal layer of actively growing shoots. The activity of the Blec4 promoter in the epidermis of the shoot apex makes it particularly suitable for genetically engineering defense against insects 10 and diseases that attack the growing shoot apex (Mandaci (1997) *Plant Mol Biol.* 34:961-965).

The invention also provides for use of tissue-specific plant promoters include a promoter from the ovule-specific *BEL1* gene described in Reiser (1995) *Cell* 83:735-742, GenBank No. U39944. Suitable seed specific promoters are derived from the 15 following genes: *MAC1* from maize, Sheridan (1996) *Genetics* 142:1009-1020; *Cat3* from maize, GenBank No. L05934, Abler (1993) *Plant Mol. Biol.* 22:10131-1038; the gene encoding oleosin 18kD from maize, GenBank No. J05212, Lee (1994) *Plant Mol. Biol.* 26:1981-1987; viviparous-1 from *Arabidopsis*, Genbank No. U93215; the gene encoding oleosin from *Arabidopsis*, Genbank No. Z17657; Atmyc1 from *Arabidopsis*, Urao (1996) 20 *Plant Mol. Biol.* 32:571-576; the 2s seed storage protein gene family from *Arabidopsis*, Conceicao (1994) *Plant* 5:493-505; the gene encoding oleosin 20kD from *Brassica napus*, GenBank No. M63985; *napA* from *Brassica napus*, GenBank No. J02798, Josefsson (1987) *JBL* 26:12196-1301; the napin gene family from *Brassica napus*, Sjodahl (1995) *Planita* 197:264-271; the gene encoding the 2S storage protein from *Brassica napus*, Dasgupta (1993) *Gene* 133:301-302; the genes encoding oleosin a, Genbank No. U09118, 25 and, oleosin B, Genbank No. U09119, from soybean; and, the gene encoding low molecular weight sulphur rich protein from soybean; Choi (1995) *Mol Gen. Genet.* 246:266-268. The tissue specific E8 promoter from tomato is particularly useful for directing gene expression so that a desired gene product is located in fruits. Other suitable 30 promoters include those from genes encoding embryonic storage proteins.

One of skill will recognize that a tissue-specific promoter may drive expression of operably linked sequences in tissues other than the target tissue. Thus, as

used herein a tissue-specific promoter is one that drives expression preferentially in the target tissue, but may also lead to some expression in other tissues as well.

The invention also provides for use of tissue-specific promoters derived from viruses which can include, *e.g.*, the tobamovirus subgenomic promoter (Kumagai (1995) *Proc. Natl. Acad. Sci. USA* 92:1679-1683; the rice tungro bacilliform virus (RTBV), which replicates only in phloem cells in infected rice plants, with its promoter which drives strong phloem-specific reporter gene expression; the cassava vein mosaic virus (CVMV) promoter, with highest activity in vascular elements, in leaf mesophyll cells, and in root tips (Verdaguer (1996) *Plant Mol. Biol.* 31:1129-1139).

In some embodiments, the nucleic acid construct will comprise a promoter functional in a specific plant cell, such as in a species of *Lactuca*, operably linked to an RG polynucleotide. Promoters useful in these embodiments include RG promoters. In additional embodiments, the nucleic acid construct will comprise a RG promoter operably linked to a heterologous polynucleotide. The heterologous polynucleotide is chosen to provide a plant with a desired phenotype. For example, the heterologous polynucleotide can be a structural gene which encodes a polypeptide which imparts a desired resistance phenotype. Alternatively, the heterologous polynucleotide may be a regulatory gene which might play a role in transcriptional and/or translational control to suppress, enhance, or otherwise modify the transcription and/or expression of an endogenous gene within the plant. The heterologous polynucleotide of the nucleic acid construct of the present invention can be expressed in either sense or anti-sense orientation as desired. It will be appreciated that control of gene expression in either sense or anti-sense orientation can have a direct impact on the observable plant characteristics.

Modifying and Inhibiting RG Gene Expression

The invention also provides for RG nucleic acid sequences which are complementary to the RG polypeptide-encoding sequences of the invention; *i.e.*, antisense RG nucleic acids. Antisense technology can be conveniently used to modify gene expression in plants. To accomplish this, a nucleic acid segment from the desired gene is cloned and operably linked to a promoter such that the anti-sense strand of RNA will be transcribed. The construct is then transformed into plants and the antisense strand of RNA is produced. In plant cells, it has been shown that antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, *see, e.g.*,

Sheehy (1988) *Proc. Nat. Acad. Sci. USA* 85:8805-8809; Hiatt et al., U.S. Patent No. 4,801,340.

Antisense sequences are capable of inhibiting the transport, splicing or transcription of RG-encoding genes. The inhibition can be effected through the targeting of genomic DNA or messenger RNA. The transcription or function of targeted nucleic acid can be inhibited, e.g., by hybridization and/or cleavage. One particularly useful set of inhibitors provided by the present invention includes oligonucleotides which are able to either bind RG gene or message, in either case preventing or inhibiting the production or function of RG. The association can be through sequence specific hybridization. Such inhibitory nucleic acid sequences can, for example, be used to completely inhibit a plant disease resistance response. Another useful class of inhibitors includes oligonucleotides which cause inactivation or cleavage of RG message. The oligonucleotide can have enzyme activity which causes such cleavage, such as ribozymes. The oligonucleotide can be chemically modified or conjugated to an enzyme or composition capable of cleaving the complementary nucleic acid. One may screen a pool of many different such oligonucleotides for those with the desired activity.

Antisense Oligonucleotides

The invention provides for antisense oligonucleotides capable of binding RG message which can inhibit RG activity by targeting mRNA. Strategies for designing antisense oligonucleotides are well described in the scientific and patent literature, and the skilled artisan can design such RG oligonucleotides using the novel reagents of the invention. In some situations, naturally occurring nucleic acids used as antisense oligonucleotides may need to be relatively long (18 to 40 nucleotides) and present at high concentrations. A wide variety of synthetic, non-naturally occurring nucleotide and nucleic acid analogues are known which can address this potential problem. For example, peptide nucleic acids (PNAs) containing non-ionic backbones, such as N-(2-aminoethyl) glycine units can be used. Antisense oligonucleotides having phosphorothioate linkages can also be used, as described in WO 97/03211; WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144:189-197; Antisense Therapeutics, ed. Agrawal (Humana Press, Totowa, N.J., 1996). Antisense oligonucleotides having synthetic DNA backbone analogues provided by the invention can also include phosphoro-dithioate, methylphosphonate,

phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, and morpholino carbamate nucleic acids, as described herein.

Combinatorial chemistry methodology can be used to create vast numbers of oligonucleotides that can be rapidly screened for specific oligonucleotides that have 5 appropriate binding affinities and specificities toward any target, such as the sense and antisense RG sequences of the invention (for general background information, see, e.g., Gold (1995) *J. of Biol. Chem.* 270:13581-13584).

Inhibitory Ribozymes

The invention provides for with ribozymes capable of binding RG message 10 which can inhibit RG activity by targeting mRNA. Strategies for designing ribozymes and selecting the RG-specific antisense sequence for targeting are well described in the scientific and patent literature, and the skilled artisan can design such RG ribozymes using the novel reagents of the invention. Ribozymes act by binding to a target RNA through the target RNA binding portion of a ribozyme which is held in close proximity to an enzymatic 15 portion of the RNA that cleaves the target RNA. Thus, the ribozyme recognizes and binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cleave and inactivate the target RNA. Cleavage of a target RNA in such a manner will destroy its ability to direct synthesis of an encoded protein if the cleavage occurs in the coding sequence, or, preventing transport of the message from the nucleus to 20 the cytoplasm. After a ribozyme has bound and cleaved its RNA target, it is typically released from that RNA and so can bind and cleave new targets repeatedly.

Catalytic RNA molecules or ribozymes can also be used to inhibit 25 expression of any plant gene. It is possible to design ribozymes that specifically pair with virtually any target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. In carrying out this cleavage, the ribozyme is not itself altered, and is thus capable of recycling and cleaving other molecules, making it a true enzyme. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs. The design and use of target RNA-specific ribozymes is described, e.g., in 30 Haseioff (1988) *Nature* 334:585-591.

In some circumstances, the enzymatic nature of a ribozyme can be advantageous over other technologies, such as antisense technology (where a nucleic acid

molecule simply binds to a nucleic acid target to block its transcription, translation or association with another molecule) as the effective concentration of ribozyme necessary to effect a therapeutic treatment can be lower than that of an antisense oligonucleotide. This potential advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single
5 ribozyme molecule is able to cleave many molecules of target RNA. In addition, a ribozyme is typically a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding, but also on the mechanism by which the molecule inhibits the expression of the RNA to which it binds. That is, the inhibition is caused by cleavage of the RNA target and so specificity is defined as the ratio
10 of the rate of cleavage of the targeted RNA over the rate of cleavage of non-targeted RNA. This cleavage mechanism is dependent upon factors additional to those involved in base pairing. Thus, the specificity of action of a ribozyme can be greater than that of antisense oligonucleotide binding the same RNA site.

The enzymatic ribozyme RNA molecule can be formed in a hammerhead
15 motif, but may also be formed in the motif of a hairpin, hepatitis delta virus, group I intron or RNaseP-like RNA (in association with an RNA guide sequence). Examples of such hammerhead motifs are described by Rossi (1992) *Aids Research and Human Retroviruses* 8:183; hairpin motifs by Hampel (1989) *Biochemistry* 28:4929, and Hampel (1990) *Nuc. Acids Res.* 18:299; the hepatitis delta virus motif by Perrotta (1992) *Biochemistry* 31:16;
20 the RNaseP motif by Guerrier-Takada (1983) *Cell* 35:849; and the group I intron by Cech U.S. Pat. No. 4,987,071. The recitation of these specific motifs is not intended to be limiting; those skilled in the art will recognize that an enzymatic RNA molecule of this invention has a specific substrate binding site complementary to one or more of the target gene RNA regions, and has nucleotide sequence within or surrounding that substrate
25 binding site which imparts an RNA cleaving activity to the molecule.

Sense Suppression

Another method of suppression is sense suppression. Introduction of
nucleic acid configured in the sense orientation has been shown to be an effective means by
which to block the transcription of target genes. For an example of the use of this method
30 to modulate expression of endogenous genes see, Napoli et al., *The Plant Cell* 2:279-289
(1990), and U.S. Patent No. 5,034,323.

Cloning of RG Polypeptides

Synthesis and/or cloning of RG polynucleotides and isolated nucleic acid constructs of the present invention are provided by methods well known to those of ordinary skill in the art. Generally, the nomenclature and the laboratory procedures in recombinant DNA technology described below are those well known and commonly employed in the art. Standard techniques are used for cloning, DNA and RNA isolation, amplification and purification. Generally enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like are performed according to the manufacturer's specifications. These techniques and various other techniques are generally performed according to Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, (1989).

The isolation of RG genes may be accomplished by a number of techniques. For instance, oligonucleotide probes based on the sequences disclosed here can be used to identify the desired gene in a cDNA or genomic DNA library. To construct genomic libraries, large segments of genomic DNA are generated by random fragmentation, e.g. using restriction endonucleases, and are ligated with vector DNA to form concatemers that can be packaged into the appropriate vector. To prepare a cDNA library, mRNA is isolated from the desired organ, such as roots and a cDNA library which contains the RG gene transcript is prepared from the mRNA. Alternatively, cDNA may be prepared from mRNA extracted from other tissues in which RG genes or homologs are expressed.

The cDNA or genomic library can then be screened using a probe based upon the sequence of a cloned RG gene such as the genes disclosed herein. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species.

Those of skill in the art will appreciate that various degrees of stringency of hybridization can be employed in the assay; and either the hybridization or the wash medium can be stringent. As the conditions for hybridization become more stringent, there must be a greater degree of complementarity between the probe and the target for duplex formation to occur. The degree of stringency can be controlled by temperature, ionic strength, pH and the presence of a partially denaturing solvent such as formamide. For example, the stringency of hybridization is conveniently varied by changing the polarity of the reactant solution through manipulation of the concentration of formamide within the range of 0% to 50%.

Alternatively, the RG nucleic acids of the invention can be amplified from nucleic acid samples using a variety of amplification techniques, such as polymerase chain reaction (PCR) technology, to amplify the sequences of the RG and related genes directly from genomic DNA, from cDNA, from genomic libraries or cDNA libraries. PCR and other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes.

Oligonucleotides can be used to identify and detect additional RG families and RG family species using a variety of hybridization techniques and conditions. Suitable amplification methods include, but are not limited to: polymerase chain reaction, PCR (PCR PROTOCOLS, A GUIDE TO METHODS AND APPLICATIONS, ed. Innis, Academic Press, N.Y. (1990) and PCR STRATEGIES (1995), ed. Innis, Academic Press, Inc., N.Y. (Innis)), ligase chain reaction (LCR) (Wu (1989) *Genomics* 4:560; Landegren (1988) *Science* 241:1077; Barringer (1990) *Gene* 89:117); transcription amplification (Kwoh (1989) *Proc. Natl. Acad. Sci. USA* 86:1173); and, self-sustained sequence replication (Guatelli (1990) *Proc. Natl. Acad. Sci. USA*, 87:1874); Q Beta replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA, Cangene, Mississauga, Ontario); see Berger (1987) *Methods Enzymol.* 152:307-316, Sambrook, and Ausubel, as well as Mullis (1987) U.S. Patent Nos. 4,683,195 and 4,683,202; Arnheim (1990) *C&EN* 36-47; Lomell *J. Clin. Chem.*, 35:1826 (1989); Van Brunt, *Biotechnology*, 8:291-294 (1990); Wu (1989) *Gene* 4:560; Sooknanan (1995) *Biotechnology* 13:563-564. Methods for cloning *in vitro* amplified nucleic acids are described in Wallace, U.S. Pat. No. 5,426,039.

The degree of complementarity (sequence identity) required for detectable binding will vary in accordance with the stringency of the hybridization medium and/or wash medium. The degree of complementarity will optimally be 100 percent; however, it should be understood that minor sequence variations in the probes and primers may be compensated for by reducing the stringency of the hybridization and/or wash medium as described earlier.

In some preferred embodiments, members of this class of pest resistance genes can be identified by their ability to be amplified by PCR primers based on the sequences disclosed here. Appropriate primers and probes for identifying RG sequences

from plant tissues are generated from comparisons of the sequences provided herein. See, e.g., Table 1. For a general overview of PCR see *PCR Protocols: A Guide to Methods and Applications*. (Innis, M, Gelfand, D., Sninsky, J. and White, T., eds.), Academic Press, San Diego (1990), incorporated herein by reference.

5 Briefly, the first step of each cycle of the PCR involves the separation of the nucleic acid duplex formed by the primer extension. Once the strands are separated, the next step in PCR involves hybridizing the separated strands with primers that flank the target sequence. The primers are then extended to form complementary copies of the target strands. For successful PCR amplification, the primers are designed so that the
10 position at which each primer hybridizes along a duplex sequence is such that an extension product synthesized from one primer, when separated from the template (complement), serves as a template for the extension of the other primer. The cycle of denaturation, hybridization, and extension is repeated as many times as necessary to obtain the desired amount of amplified nucleic acid.

15 In the preferred embodiment of the PCR process, strand separation is achieved by heating the reaction to a sufficiently high temperature for an sufficient time to cause the denaturation of the duplex but not to cause an irreversible denaturation of the polymerase (see U.S. Patent No. 4,965,188). Template-dependent extension of primers in PCR is catalyzed by a polymerizing agent in the presence of adequate amounts of four
20 deoxyribonucleotide triphosphates (typically dATP, dGTP, dCTP, and dTTP) in a reaction medium comprised of the appropriate salts, metal cations, and pH buffering system. Suitable polymerizing agents are enzymes known to catalyze template-dependent DNA synthesis.

25 Polynucleotides may also be synthesized by well-known techniques as described in the technical literature. See, e.g., Carruthers *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 47:411-418 (1982), and Adams *et al.*, *J. Am. Chem. Soc.* 105:661 (1983). Double stranded DNA fragments may then be obtained either by synthesizing the complementary strand and annealing the strands together under appropriate conditions, or by adding the complementary strand using DNA polymerase with an appropriate primer
30 sequence.

The present invention further provides isolated RG proteins encoded by the RG polynucleotides disclosed herein. One of skill will recognize that the nucleic acid encoding a functional RG protein need not have a sequence identical to the exemplified genes disclosed here. For example, because of codon degeneracy a large number of 5 nucleic acid sequences can encode the same polypeptide. In addition, the polypeptides encoded by the RG genes, like other proteins, have different domains which perform different functions. Thus, the RG gene sequences need not be full length, so long as the desired functional domain of the protein is expressed.

The resistance proteins are at least 25 amino acid residues in length. 10 Typically, the RG proteins are at least 50 amino acid residues, generally at least 100, preferably at least 150, more preferably at least 200 amino acids in length. In particularly preferred embodiments, the RG proteins are of sufficient length to provide resistance to pests when expressed in the desired plants. Generally then, the RG proteins will be the length encoded by an RG gene of the present invention. However, those of ordinary skill 15 will appreciate that minor deletions, substitutions, or additions to an RG protein will typically yield a protein with pest resistance characteristics similar or identical to that of the full length sequence. Thus, full-length RG proteins modified by 1, 2, 3, 4, or 5 deletions, substitutions, or additions, generally provide an effective degree of pest resistance relative to the full-length protein.

20 The RG proteins which provide pest resistance will typically comprise at least one of an LRR or an NBS. Preferably, both are present. LRR and/or NBS regions present in the RG proteins of the present invention can be provided by RG genes of the present invention. In some embodiments, the LRR and/or NBS regions are obtained from other pest resistance genes. See, e.g., Yu *et al.*, *Proc. Natl. Acad. Sci. USA*, 93: 11751-25 11756 (1996); Bent *et al.*, *Science*, 265: 1856-1860 (1994).

Modified protein chains can also be readily designed utilizing various recombinant DNA techniques well known to those skilled in the art. For example, the chains can vary from the naturally occurring sequence at the primary structure level by amino acid substitutions, additions, deletions, and the like. Modification can also include 30 swapping domains from the proteins of the invention with related domains from other pest resistance genes.

Pests that can be targeted by RG genes and proteins of the present invention include such bacterial pests as *Erwinia carotovora* and *Pseudomonas marginalis*. Fungal pests which can be targeted by the present invention include *Bremia lactucae*, *Marssonina panattoniana*, *Rhizoctonia solani*, *Olpidium brassicae*, root aphid, *Sclerotinia sclerotiorum* and *S. minor*, and *Botrytis cinerea* which causes gray mold. RG genes also provide resistance to viral diseases such as lettuce and turnip mosaic viruses.

5 **Fusion Proteins**

RG polypeptides can also be expressed as recombinant proteins with one or more additional polypeptide domains linked thereto to facilitate protein detection, 10 purification, or other applications. Such detection and purification facilitating domains include, but are not limited to, metal chelating peptides such as polyhistidine tracts and histidine-tryptophan modules that allow purification on immobilized metals, protein a domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle WA). The 15 inclusion of a cleavable linker sequences such as Factor Xa or enterokinase (Invitrogen, San Diego CA) between the purification domain and plant disease resistant polypeptide may be useful to facilitate purification. One such expression vector provides for expression of a fusion protein comprising the sequence encoding a plant disease resistant polypeptide of the invention and nucleic acid sequence encoding six histidine residues 20 followed by thioredoxin and an enterokinase cleavage site (e.g., see Williams (1995) *Biochemistry* 34:1787-1797). The histidine residues facilitate detection and purification while the enterokinase cleavage site provides a means for purifying the desired protein(s) from the remainder of the fusion protein. Technology pertaining to vectors encoding 25 fusion proteins and application of fusion proteins are well described, see e.g., Kroll (1993) *DNA Cell. Biol.*, 12:441-53.

Antibodies Reactive to RG Polypeptides and Immunological Assays

The present invention also provides antibodies which specifically react with RG proteins of the present invention under immunologically reactive conditions. An antibody immunologically reactive with a particular antigen can be generated *in vivo* or by 30 recombinant methods such as selection of libraries of recombinant antibodies in phage or similar vectors. "Immunologically reactive conditions" includes reference to conditions which allow an antibody, generated to a particular epitope of an antigen, to bind to that

epitope to a detectably greater degree than the antibody binds to substantially all other epitopes, generally at least two times above background binding, preferably at least five times above background. Immunologically reactive conditions are dependent upon the format of the antibody binding reaction and typically are those utilized in immunoassay protocols.

"Antibody" includes reference to an immunoglobulin molecule obtained by *in vitro* or *in vivo* generation of the humoral response, and includes both polyclonal and monoclonal antibodies. The term also includes genetically engineered forms such as chimeric antibodies (e.g., humanized murine antibodies), heteroconjugate antibodies (e.g., bispecific antibodies), and recombinant single chain Fv fragments (scFv). The term "antibody" also includes antigen binding forms of antibodies (e.g., Fab', F(ab')₂, Fab, Fv, rIgG, and, inverted IgG). See, Pierce Catalog and Handbook, 1994-1995 (Pierce Chemical Co., Rockford, IL). An antibody immunologically reactive with a particular antigen can be generated *in vivo* or by recombinant methods such as selection of libraries of recombinant antibodies in phage or similar vectors. *See, e.g., Huse et al. (1989) Science 246:1275-1281; and Ward, et al. (1989) Nature 341:544-546; and Vaughan et al. (1996) Nature Biotechnology, 14:309-314.*

Many methods of making antibodies are known to persons of skill. A number of immunogens are used to produce antibodies specifically reactive to an isolated RG protein of the present invention under immunologically reactive conditions. An isolated recombinant, synthetic, or native RG protein of the present invention is the preferred immunogens (antigen) for the production of monoclonal or polyclonal antibodies.

The RG protein is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies can be generated for subsequent use in immunoassays to measure the presence and quantity of the RG protein. Methods of producing monoclonal or polyclonal antibodies are known to those of skill in the art. *See, e.g., Coligan (1991) Current Protocols in Immunology Wiley/Greene, NY; and Harlow and Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press, NY; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.) Academic Press, New York, NY.*

Frequently, the RG proteins and antibodies will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide

variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionucleotides, enzymes, substrates, cofactors, inhibitors, fluorescent moieties, chemiluminescent moieties, magnetic particles, and the like. Patents teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

The antibodies of the present invention can be used to screen plants for the expression of RG proteins of the present invention. The antibodies of this invention are also used for affinity chromatography in isolating RG protein.

The present invention further provides RG polypeptides that specifically bind, under immunologically reactive conditions, to an antibody generated against a defined immunogen, such as an immunogen consisting of the RG polypeptides of the present invention. Immunogens will generally be at least 10 contiguous amino acids from an RG polypeptide of the present invention. Optionally, immunogens can be from regions exclusive of the NBS and/or LRR regions of the RG polypeptides. Nucleic acids which encode such cross-reactive RG polypeptides are also provided by the present invention. The RG polypeptides can be isolated from any number plants as discussed earlier. Preferred are species from the family *Compositae* and in particular the genus *Lactuca* such as *L. sativa* and such subspecies as *crispa*, *longifolia*, and *asparagina*.

"Specifically binds" includes reference to the preferential association of a ligand, in whole or part, with a particular target molecule (i.e., "binding partner" or "binding moiety") relative to compositions lacking that target molecule. It is, of course, recognized that a certain degree of non-specific interaction may occur between a ligand and a non-target molecule. Nevertheless, specific binding, may be distinguished as mediated through specific recognition of the target molecule. Typically specific binding results in a much stronger association between the ligand and the target molecule than between the ligand and non-target molecule. Specific binding by an antibody to a protein under such conditions requires an antibody that is selected for its specificity for a particular protein. The affinity constant of the antibody binding site for its cognate monovalent antigen is at least 10^7 , usually at least 10^8 , preferably at least 10^9 , more preferably at least 10^{10} , and most preferably at least 10^{11} liters/mole. A variety of immunoassay formats are appropriate for selecting antibodies specifically reactive with a particular protein. For

example, solid-phase ELISA immunoassays are routinely used to select monoclonal antibodies specifically reactive with a protein. See Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York, for a description of immunoassay formats and conditions that can be used to determine specific reactivity. The 5 antibody may be polyclonal but preferably is monoclonal. Generally, antibodies cross-reactive to such proteins as RPS2, RPM1 (bacterial resistances in *Arabidopsis*, L6 (fungal resistance in flax, PRF (resistance to *Pseudomonas syringae* in tomato), and N, (virus resistance in tobacco), are removed by immunoabsorption.

Immunoassays in the competitive binding format are typically used for 10 cross-reactivity determinations. For example, an immunogenic RG polypeptide is immobilized to a solid support. Polypeptides added to the assay compete with the binding of the antisera to the immobilized antigen. The ability of the above polypeptides to compete with the binding of the antisera to the immobilized RG polypeptide is compared to the immunogenic RG polypeptide. The percent cross-reactivity for the above proteins is 15 calculated, using standard calculations. Those antisera with less than 10% cross-reactivity with such proteins as RPS2, RPM1, L6, PRF, and N, are selected and pooled. The cross-reacting antibodies are then removed from the pooled antisera by immunoabsorption with these non-RG resistance proteins.

The immunoabsorbed and pooled antisera are then used in a competitive 20 binding immunoassay to compare a second "target" polypeptide to the immunogenic polypeptide. In order to make this comparison, the two polypeptides are each assayed at a wide range of concentrations and the amount of each polypeptide required to inhibit 50% of the binding of the antisera to the immobilized protein is determined using standard techniques. If the amount of the target polypeptide required is less than twice the amount 25 of the immunogenic polypeptide that is required, then the target polypeptide is said to specifically bind to an antibody generated to the immunogenic protein. As a final determination of specificity, the pooled antisera is fully immunosorbed with the immunogenic polypeptide until no binding to the polypeptide used in the immunosorption is detectable. The fully immunosorbed antisera is then tested for reactivity with the test 30 polypeptide. If no reactivity is observed, then the test polypeptide is specifically bound by the antisera elicited by the immunogenic protein.

Production of transgenic plants of the invention

Isolated nucleic acid constructs prepared as described herein can be introduced into plants according techniques known in the art. In some embodiments, the introduced nucleic acid is used to provide RG gene expression and therefore pest resistance in desired plants. In some embodiments, RG promoters are used to drive expression of desired heterologous genes in plants. Finally, in some embodiments, the constructs can be used to suppress expression of a target endogenous gene, including RG genes.

To use isolated RG sequences in the above techniques, recombinant DNA vectors suitable for transformation of plant cells are prepared. Techniques for transforming a wide variety of higher plant species are well known and described in the technical and scientific literature. See, for example, Weising *et al. Ann. Rev. Genet.* 22:421-477 (1988).

A DNA sequence coding for the desired RG polypeptide, for example a cDNA or a genomic sequence encoding a full length protein, will be used to construct a recombinant expression cassette which can be introduced into the desired plant. An expression cassette will typically comprise the RG polynucleotide operably linked to transcriptional and translational initiation regulatory sequences which will direct the transcription of the sequence from the RG gene in the intended tissues of the transformed plant.

Such DNA constructs may be introduced into the genome of the desired plant host by a variety of conventional techniques. For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation, PEG poration, particle bombardment and microinjection of plant cell protoplasts or embryogenic callus, or the DNA constructs can be introduced directly to plant tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria.

Transformation techniques are known in the art and well described in the scientific and patent literature. The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski *et al. Embo J.* 3:2717-2722 (1984).

Electroporation techniques are described in Fromm *et al.* *Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Ballistic transformation techniques are described in Klein *et al.* *Nature* 327:70-73 (1987).

Agrobacterium tumefaciens-mediated transformation techniques are well described in the scientific literature. See, for example Horsch *et al.* *Science* 233:496-498 (1984), and Fraley *et al.* *Proc. Natl. Acad. Sci. USA* 80:4803 (1983). Although Agrobacterium is useful primarily in dicots, certain monocots can be transformed by Agrobacterium. For instance, Agrobacterium transformation of rice is described by Hiei *et al.*, *Plant J.* 6:271-282 (1994). A particularly preferred means of transforming lettuce is described in Michelmore *et al.*, *Plant Cell Reports*, 6:439-442 (1987).

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant which possesses the transformed genotype and thus the desired RG-controlled phenotype. Such regeneration techniques rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with the RG nucleotide sequences. Plant regeneration from cultured protoplasts is described in Evans *et al.*, *Protoplasts Isolation and Culture, Handbook of Plant Cell Culture*, pp. 124-176, Macmillan Publishing Company, New York, 1983; and Binding, *Regeneration of Plants, Plant Protoplasts*, pp. 21-73, CRC Press, Boca Raton, 1985. Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee *et al.* *Ann. Rev. of Plant Phys.* 38:467-486 (1987).

The methods of the present invention are particularly useful for incorporating the RG polynucleotides into transformed plants in ways and under circumstances which are not found naturally. In particular, the RG polypeptides may be expressed at times or in quantities which are not characteristic of natural plants.

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

Detection of RG Resistance Genes

The present invention further provides methods for detecting RG resistance genes in a nucleic acid sample suspected of comprising an RG resistance gene. The means by which the RG resistance gene is detected is not a critical aspect of the invention. For example, RG resistance genes can be detected by the presence of amplicons using RG resistance gene specific primers. Additionally, RG resistance genes can be detected by assaying for specific hybridization of an RG polynucleotide to an RG resistance gene. In some embodiments, the RG resistance gene can be amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide.

In a typical detection method, the nucleic acid sample is contacted with an RG polynucleotide to form a hybridization complex. The hybridization complex may be detected directly (e.g., in Southern or northern blots), or indirectly (e.g., by subsequent primer extension during PCR amplification). The RG polynucleotide hybridizes under stringent conditions to an RG polynucleotide of the invention. Formation of the hybridization complex is directly or indirectly used to indicate the presence of the RG resistance gene in the nucleic acid sample.

Detection of the hybridization complex can be achieved using any number of well known methods. For example, the nucleic acid sample, or a portion thereof, may be assayed by hybridization formats including but not limited to, solution phase, solid phase, mixed phase, or *in situ* hybridization assays. Briefly, in solution (or liquid) phase hybridizations, both the target nucleic acid and the probe or primer are free to interact in the reaction mixture. In solid phase hybridization assays, probes or primers are typically linked to a solid support where they are available for hybridization with target nucleic acid in solution. In mixed phase, nucleic acid intermediates in solution hybridize to target nucleic acids in solution as well as to a nucleic acid linked to a solid support. In *in situ* hybridization, the target nucleic acid is liberated from its cellular surroundings in such as to be available for hybridization within the cell while preserving the cellular morphology for subsequent interpretation and analysis. The following articles provide an overview of the various hybridization assay formats: Singer *et al.*, *Biotechniques* 4(3):230-250 (1986); Haase *et al.*, *Methods in Virology*, Vol. VII, pp. 189-226 (1984); Wilkinson, "The theory and practice of *in situ* hybridization" In: *In situ Hybridization*, Ed. D.G. Wilkinson. IRL Press, Oxford University Press, Oxford; and *Nucleic Acid Hybridization: A Practical Approach*, Ed. Hames, B.D. and Higgins, S.J., IRL Press (1987).

The effect of the modification of RG gene expression can be measured by detection of increases or decreases in mRNA levels using, for instance, Northern blots. In addition, the phenotypic effects of gene expression can be detected by measuring nematode, fungal, bacterial, viral, or other pest resistance in plants. Suitable assays for determining pest resistance are well known. Michelmore and Crute, *Trans. Br. mycol. Soc.*, 79(3): 542-546 (1982).

The means by which hybridization complexes are detected is not a critical aspect of the present invention and can be accomplished by any number of methods currently known or later developed. RG polynucleotides can be labeled by any one of several methods typically used to detect the presence of hybridized nucleic acids. One common method of detection is the use of autoradiography using probes labeled with ³H, ¹²⁵I, ³⁵S, ¹⁴C, or ³²P, or the like. The choice of radioactive isotope depends on research preferences due to ease of synthesis, stability, and half lives of the selected isotopes. Other labels include ligands which bind to antibodies labeled with fluorophores, chemiluminescent agents, and enzymes. Alternatively, probes can be conjugated directly with labels such as fluorophores, chemiluminescent agents or enzymes. The choice of label depends on sensitivity required, ease of conjugation with the probe, stability requirements, and available instrumentation. Labeling the RG polynucleotide is readily achieved such as by the use of labeled PCR primers.

The choice of label dictates the manner in which the label is bound to the probe. Radioactive probes are typically made using commercially available nucleotides containing the desired radioactive isotope. The radioactive nucleotides can be incorporated into probes, for example, by using DNA synthesizers, by nick translation with DNA polymerase I, by tailing radioactive DNA bases to the 3' end of probes with terminal deoxynucleotidyl transferase, by treating single-stranded M13 plasmids having specific inserts with the Klenow fragment of DNA polymerase in the presence of radioactive deoxynucleotides, dNTP, by transcribing from RNA templates using reverse transcriptase in the presence of radioactive deoxynucleotides, dNTP, or by transcribing RNA from vectors containing specific RNA viral promoters (e.g., SP6 promoter) using the corresponding RNA polymerase (e.g., SP6 RNA polymerase) in the presence of radioactive ribonucleotides rNTP.

The probes can be labeled using radioactive nucleotides in which the isotope resides as a part of the nucleotide molecule, or in which the radioactive component is attached to the nucleotide via a terminal hydroxyl group that has been esterified to a radioactive component such as inorganic acids, e.g., ³²P phosphate or ¹⁴C organic acids, or esterified to provide a linking group to the label. Base analogs having nucleophilic linking groups, such as primary amino groups, can also be linked to a label.

Non-radioactive probes are often labeled by indirect means. For example, a ligand molecule is covalently bound to the probe. The ligand then binds to an anti-ligand molecule which is either inherently detectable or covalently bound to a detectable signal system, such as an enzyme, a fluorophore, or a chemiluminescent compound. Enzymes of interest as labels will primarily be hydrolases, such as phosphatases, esterases and glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescers include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. Ligands and anti-ligands may be varied widely. Where a ligand has a natural anti-ligand, namely ligands such as biotin, thyroxine, and cortisol, it can be used in conjunction with its labeled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

Probes can also be labeled by direct conjugation with a label. For example, cloned DNA probes have been coupled directly to horseradish peroxidase or alkaline phosphatase, (Renz, M., and Kurz, K. (1984) A Colorimetric Method for DNA Hybridization. *Nucl. Acids Res.* 12: 3435-3444) and synthetic oligonucleotides have been coupled directly with alkaline phosphatase (Jablonski, E., et al. (1986) Preparation of Oligodeoxynucleotide-Alkaline Phosphatase Conjugates and Their Use as Hybridization Probes. *Nuc. Acids. Res.* 14: 6115-6128; and Li P., et al. (1987) Enzyme-linked Synthetic Oligonucleotide probes: Non-Radioactive Detection of Enterotoxigenic *Escherichia Coli* in Faecal Specimens. *Nucl. Acids Res.* 15:5275-5287).

Definitions

Units, prefixes, and symbols can be denoted in their SI accepted form. Numeric ranges are inclusive of the numbers defining the range. Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation, respectively. The

headings provided herein are not limitations of the various aspects or embodiments of the invention which can be had by reference to the specification as a whole. Accordingly, the terms defined immediately below are more fully defined by reference to the specification as a whole.

5 As used herein, the term "plant" includes reference to whole plants, plant organs (e.g., leaves, stems, roots, etc.), seeds and plant cells and progeny of same. The class of plants which can be used in the methods of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants.

10 As used herein, "pest" includes, but is not limited to, viruses, fungi, nematodes, insects, and bacteria.

15 As used herein, "heterologous" is a nucleic acid that originates from a foreign species, or, if from the same species, is substantially modified from its original form. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived, or, if from the same species, one or both are substantially modified from their original form.

20 As used herein, "RG gene," alternatively referred to as "RLG gene," is a gene encoding resistance to plant pests, such as viruses, fungi, nematodes, insects, and bacteria, and which hybridizes under stringent conditions and/or has at least 60% sequence identity at the deduced amino acid level to the exemplified sequences provided herein. RG genes encode "RG polypeptides," alternatively referred to as "RLG polypeptides," which can comprise LRR motifs and/or NBS motifs. The RG polypeptides encoded by RG genes have at least 55% or 60% sequence identity, typically at least 65% sequence identity, preferably at least 70% sequence identity, often at least 75% sequence identity, more 25 preferably at least 80% sequence identity, and most preferably at least 90% sequence identity at the deduced amino acid level relative to the exemplary RG sequences provided herein. The term "RG family" or "RG family genus" or "genus" includes reference to a group of RG polypeptide sequence species that have at least 60% amino acid sequence identity, and, the nucleic acids encoding these polypeptides. The individual species of a genus, i.e., the members of a family, typically are genetically mapped to the same locus.

30 As used herein, "RG polynucleotide" includes reference to a contiguous sequence from an RG gene of at least 18, 20, 25, 30, 40, or 50 nucleotides in length, up to

at least about 100 or at least about 200 nucleotides in length. In some embodiments, the polynucleotide is preferably at least 100 nucleotides in length, more preferably at least 200 nucleotides in length, most preferably at least 500 nucleotides in length. Thus, RG polynucleotide may be a RG gene or a subsequence thereof.

5 As used herein, "isolated," when referring to a molecule or composition, such as, for example, an RG polypeptide or nucleic acid, means that the molecule or composition is separated from at least one other compound, such as a protein, other nucleic acids (*e.g.*, RNAs), or other contaminants with which it is associated *in vivo* or in its naturally occurring state. Thus, an RG polypeptide or nucleic acid is considered isolated
10 when it has been isolated from any other component with which it is naturally associated, *e.g.*, cell membrane, as in a cell extract. An isolated composition can, however, also be substantially pure. An isolated composition can be in a homogeneous state and can be in a dry or an aqueous solution. Purity and homogeneity can be determined, for example, using analytical chemistry techniques such as polyacrylamide gel electrophoresis (SDS-
15 PAGE) or high performance liquid chromatography (HPLC).

The term "nucleic acid" or "nucleic acid molecule" or "nucleic acid sequence" refers to a deoxyribonucleotide or ribonucleotide oligonucleotide in either single- or double-stranded form. The term encompasses nucleic acids, *i.e.*, oligonucleotides, containing known analogues of natural nucleotides which have similar or
20 improved binding properties, for the purposes desired, as the reference nucleic acid. The term also includes nucleic acids which are metabolized in a manner similar to naturally occurring nucleotides or at rates that are improved thereover for the purposes desired. The term also encompasses nucleic-acid-like structures with synthetic backbones. DNA backbone analogues provided by the invention include phosphodiester, phosphorothioate,
25 phosphorodithioate, methylphosphonate, phosphoramidate, alkyl phosphotriester, sulfamate, 3'-thioacetal, methylene(methylimino), 3'-N-carbamate, morpholino carbamate, and peptide nucleic acids (PNAs); see Oligonucleotides and Analogues, a Practical Approach, edited by F. Eckstein, IRL Press at Oxford University Press (1991); Antisense Strategies, Annals of the New York Academy of Sciences, Volume 600, Eds. Baserga and
30 Denhardt (NYAS 1992); Milligan (1993) *J. Med. Chem.* 36:1923-1937; Antisense Research and Applications (1993, CRC Press). PNAs contain non-ionic backbones, such as N-(2-aminoethyl) glycine units. Phosphorothioate linkages are described in WO

97/03211; WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144:189-197. Other synthetic backbones encompassed by the term include methyl-phosphonate linkages or alternating methylphosphonate and phosphodiester linkages (Strauss-Soukup (1997) *Biochemistry* 36:8692-8698), and benzylphosphonate linkages (Samstag (1996) *Antisense Nucleic Acid Drug Dev* 6:153-156). The term nucleic acid is used interchangeably with gene, cDNA, mRNA, oligonucleotide primer, probe and amplification product. Unless otherwise indicated, a particular nucleic acid sequence includes the complementary sequence thereof.

The term "exogenous nucleic acid" refers to a nucleic acid that has been isolated, synthesized, cloned, ligated, excised in conjunction with another nucleic acid, in a manner that is not found in nature, and/or introduced into and/or expressed in a cell or cellular environment other than or at levels or forms different than the cell or cellular environment in which said nucleic acid or protein is found in nature. The term encompasses both nucleic acids originally obtained from a different organism or cell type than the cell type in which it is expressed, and also nucleic acids that are obtained from the same cell line as the cell line in which it is expressed. invention.

The term "recombinant," when used with reference to a cell, or to the nucleic acid, protein or vector refers to a material, or a material corresponding to the natural or native form of the material, that has been modified by the introduction of a new moiety or alteration of an existing moiety, or is identical thereto but produced or derived from synthetic materials. For example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise expressed at a different level, typically, under-expressed or not expressed at all. The term "recombinant means" encompasses all means of expressing, i.e., transcription or translation of, an isolated and/or cloned nucleic acid *in vitro* or *in vivo*. For example, the term "recombinant means" encompasses techniques where a recombinant nucleic acid, such as a cDNA encoding a protein, is inserted into an expression vector, the vector is introduced into a cell and the cell expresses the protein. "Recombinant means" also encompass the ligation of nucleic acids having coding or promoter sequences from different sources into one vector for expression of a fusion protein, constitutive expression of a protein, or inducible expression of a protein, such as the plant disease resistant, or RG. polypeptides of the invention.

The term "specifically hybridizes" refers to a nucleic acid that hybridizes, duplexes or binds to a particular target DNA or RNA sequence. The target sequences can be present in a preparation of total cellular DNA or RNA. Proper annealing conditions depend, for example, upon a nucleic acid's, such as a probe's length, base composition, 5 and the number of mismatches and their position on the probe, and can be readily determined empirically providing the appropriate reagents are available. For discussions of nucleic acid probe design and annealing conditions, see, e.g., Sambrook and Ausubel.

The terms "stringent hybridization," "stringent conditions," or "specific hybridization conditions" refers to conditions under which an oligonucleotide (when used, 10 for example, as a probe or primer) will hybridize to its target subsequence, such as an RG nucleic acid in an expression vector of the invention but not to a non-RG sequence. Stringent conditions are sequence-dependent. Thus, in one set of stringent conditions an oligonucleotide probe will hybridize to only one species of the genus of RG nucleic acids of the invention. In another set of stringent conditions (less stringent) an oligonucleotide 15 probe will hybridize to all species of the invention's genus but not to non-RG nucleic acids. Longer sequences hybridize specifically at higher temperatures. Stringent conditions are selected to be about 5⁰C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes 20 complementary to the target sequence hybridize to the target sequence at equilibrium (if the target sequences are present in excess, at T_m , 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, i.e., about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30⁰C for short probes 25 (e.g., 10 to 50 nucleotides) and at least about 60⁰C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Often, high stringency wash conditions preceded by low 30 stringency wash conditions to remove background probe signal. An example of medium stringency wash conditions for a duplex of, e.g., more than 100 nucleotides, is 1x SSC at 45⁰C for 15 minutes (see Sambrook for a description of SSC buffer). An example low stringency wash for a duplex of, e.g., more than 100 nucleotides, is 4-6x SSC at 40⁰C for 15 minutes. a signal to noise ratio of 2x (or higher) than that observed for an unrelated

probe in the particular hybridization assay indicates detection of a "specific hybridization." Nucleic acids which do not hybridize to each other under stringent conditions can still be substantially identical if the polypeptides which they encode are substantially identical. This can occurs, *e.g.*, when a nucleic acid is created that encodes for conservative substitutions. Stringent hybridization and stringent hybridization wash conditions are different under different environmental parameters, such as for Southern and Northern hybridizations. An extensive guide to the hybridization of nucleic acids is found in, *e.g.*, Sambrook, Tijssen (1993) *supra*.

As used herein "operably linked" includes reference to a functional linkage between a promoter and a second sequence, wherein the promoter sequence initiates and mediates transcription of the DNA sequence corresponding to the second sequence. Generally, operably linked means that the nucleic acid sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in the same reading frame.

In the expression of transgenes one of skill will recognize that the inserted polynucleotide sequence need not be identical and may be "substantially identical" to a sequence of the gene from which it was derived. As explained herein, these variants are specifically covered by this term.

In the case where the inserted polynucleotide sequence is transcribed and translated to produce a functional RG polypeptide, one of skill will recognize that because of codon degeneracy, a number of polynucleotide sequences will encode the same polypeptide. These variants are specifically covered by the term "RG polynucleotide sequence". In addition, the term specifically includes those full length sequences substantially identical (determined as described herein) with an RG gene sequence which encode proteins that retain the function of the RG protein. Thus, in the case of RG genes disclosed here, the term includes variant polynucleotide sequences which have substantial identity with the sequences disclosed here and which encode proteins capable of conferring resistance to nematodes, bacteria, viruses, fungi, insects or other pests on a transgenic plant comprising the sequence.

Two polynucleotides or polypeptides are said to be "identical" if the sequence of nucleotides or amino acid residues, respectively, in the two sequences is the same when aligned for maximum correspondence, as described below. The term

"complementary to" is used herein to mean that the complementary sequence is identical to all or a specified contiguous portion of a reference polynucleotide sequence.

The terms "sequence identity," "sequence similarity" and "homology" refer to when two sequences, such as the nucleic acid and amino acid sequences or the 5 polypeptides of the invention, when optimally aligned, as with, for example, the programs PILEUP, BLAST, GAP, FASTA or BESTFIT (see discussion, *supra*). "Percentage amino acid/nucleic acid sequence identity" refers to a comparison of the sequences of two polypeptides/nucleic acids which, when optimally aligned, have approximately the designated percentage of the same amino acids/nucleic acids, respectively. For example, 10 "60% sequence identity" and "60% homology" refer to a comparison of the sequences of two RG nucleic acids or polypeptides which, when optimally aligned, have 60% identity. For example, in one embodiment, nucleic acids encoding RG polypeptides of the invention comprise a sequence with at least 50% nucleic acid sequence identity to SEQ ID NO:1. In other embodiments, the RG polypeptides of the invention are encoded by nucleic acids 15 comprising a sequence with at least 50% sequence identity to SEQ ID NO:1, or, are encoded by nucleic acids comprising SEQ ID NO:1, or, have at least 60% amino acid sequence identity to the polypeptide of SEQ ID NO:2.

"Percentage of sequence identity" is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide 20 sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of 25 matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

The term "substantial identity" of polynucleotide sequences means that a polynucleotide comprises a sequence that has at least 55% or 60% sequence identity, generally at least 65%, preferably at least 70%, often at least 75%, more preferably at 30 least 80% and most preferably at least 90%, compared to a reference sequence using the programs described above (preferably BESTFIT) using standard parameters. One of skill will recognize that these values can be appropriately adjusted to determine corresponding

identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like. Substantial identity of amino acid sequences for these purposes normally means sequence identity of at least 55% or 60%, preferably at least 70%, more preferably at least 80%, and most preferably at least 95%. Polypeptides having "sequence similarity" share sequences as noted above except that residue positions which are not identical may differ by conservative amino acid changes. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine.

Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

Another indication that nucleotide sequences are substantially identical is if two molecules hybridize to each other under appropriate conditions. Appropriate conditions can be high or low stringency and will be different in different circumstances. Generally, stringent conditions are selected to be about 5°C to about 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent wash conditions are those in which the salt concentration is about 0.02 molar at pH 7 and the temperature is at least about 50°C. However, nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This may occur, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

Nucleic acids of the invention can be identified from a cDNA or genomic library prepared according to standard procedures and the nucleic acids disclosed here used as a probe. Thus, for example, stringent hybridization conditions will typically include at least one low stringency wash using 0.3 molar salt (e.g., 2X SSC) at 65°C. The washes

are preferably followed by one or more subsequent washes using 0.03 molar salt (e.g., 0.2X SSC) at 50°C, usually 60°C, or more usually 65°C. Nucleic acid probes used to identify the nucleic acids are preferably at least 100 nucleotides in length.

As used herein, "nucleotide binding site" or "nucleotide binding domain" ("NBS") includes reference to highly conserved nucleotide-, *i.e.*, ATP/GTP-, binding domains, typically included in the "kinase domain" of kinase polypeptides, such as a kinase-1a, kinase 2, or a kinase 3a motif, as described herein. For example, the tobacco N and Arabidopsis RPS2 genes, among several recently cloned disease-resistance genes, share highly conserved NBS sequence. Kinase NBS subdomains further consist of three subdomain motifs: the P-loop, kinase-2, and kinase-3a subdomains (Yu (1996) *Proc. Acad. Sci. USA* 93:11751-11756). As discussed in detail herein, examples include the *Arabidopsis* RPP5 gene (Parker (1997) *supra*), the *A. thaliana* RPS2 gene (Mindrinos (1997) *supra*), and the flax L6 rust resistance gene (Lawrence (1995) *supra*) which all encode proteins containing an NBS; and Mindrinos (1994) *Cell* 78:1089-1099; and Shen (1993) *FEBS* 335:380-385. Using the teachings disclosed and incorporated herein and standard nucleic acid hybridization and/or amplification techniques, one of skill can identify members having NBS domains, including any of the genus of NBS-containing plant disease resistant polypeptides of the invention.

As used herein, "leucine rich region" ("LRR") includes reference to a region that has a leucine content of at least 20% leucine or isoleucine, or 30% of the aliphatic residues: leucine, isoleucine, methionine, valine, and phenylalanine, and arranged with approximate repeated periodicity. The length of the repeat may vary in length but is generally about 20 to 30 amino acids. An LRR-containing polypeptide typically will have the canonical 24 amino acid leucine-rich repeat (LRR) sequence, which is present in different proteins that mediate molecular recognition and/or interaction processes; as described in Bent (1994) *Science* 265:1856-1860; Parker (1997) *Plant Cell* 9:879-894; Hong (1997) *Plant Physiol.* 113:1203-1212; Schmitz (1997) *Nucleic Acids Res.* 25:756-763; Hipskind (1996) *Mol. Plant Microbe Interact.* 9:819-825; Tornero (1996) *Plant J.* 10:315-330; Dixon (1996) *Cell* 84:451-459; Jones (1994) *Science* 266:789-793; Lawrence (1995) *Plant Cell* 7:1195-1206; Song (1995) *Science* 270:1804-1806; as discussed in further detail *supra*. Using the teachings disclosed and incorporated herein and standard nucleic acid hybridization and/or amplification techniques, one of skill can

identify polypeptides having LRR domains, including any member of the genus of LRR-containing RG polypeptides of the invention.

The term "promoter" refers to a region or sequence determinants located upstream or downstream from the start of transcription and which are involved in 5 recognition and binding of RNA polymerase and other proteins to initiate transcription. A "plant promoter" is a promoter capable of initiating and/or regulating transcription in plant cells; see also discussion on plant promoters, *supra*.

The term "constitutive promoter" refers to a promoter that initiates and helps control transcription in all tissues. Promoters that drive expression continuously 10 under physiological conditions are referred to herein as "constitutive" promoters and are active under most environmental conditions and states of development or cell differentiation; see also detailed discussion, *supra*.

The term "inducible promoter" refers to a promoter which directs transcription under the influence of changing environmental conditions or developmental 15 conditions. Examples of environmental conditions that may effect transcription by inducible promoters include anaerobic conditions, elevated temperature, drought, or the presence of light. Such promoters are referred to herein as "inducible" promoters; see also detailed discussion, *supra*.

The term "abscission-induced promoter" or "abscission promoter" refers to a 20 class of promoters which are activated upon plant ripening, such as fruit ripening, and are especially useful incorporated in the expression systems (*e.g.*, expression cassettes, vectors) of the invention. When the plant disease resistant polypeptide-encoding nucleic acid is under the control of an abscission promoter, rapid cell death, induced by expression 25 of the invention's polypeptide, accelerates and/or accentuates abscission of the plant part, increasing the efficiency of the harvesting of fruits or other plant parts, such as cotton, and the like; see also detailed discussion, *supra*.

The term "tissue-specific promoter" refers to a class of transcriptional control elements that are only active in particular cells or tissues. Examples of plant 30 promoters under developmental control include promoters that initiate transcription only (or primarily only) in certain tissues, such as roots, leaves, fruit, ovules, seeds, pollen, pistils, or flowers; see also detailed discussion, *supra*.

As used herein "recombinant" includes reference to a cell, or nucleic acid, or vector, that has been modified by the introduction of a heterologous nucleic acid or the alteration of a native nucleic acid to a form not native to that cell, or that the cell is derived from a cell so modified. Thus, for example, recombinant cells express genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise abnormally expressed, under expressed or not expressed at all.

As used herein, a "recombinant expression cassette" or "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements which permit transcription of a particular nucleic acid in a target cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the recombinant expression cassette portion of the expression vector includes a nucleic acid to be transcribed, and a promoter.

As used herein, "transgenic plant" includes reference to a plant modified by introduction of a heterologous polynucleotide. Generally, the heterologous polynucleotide is an RG structural or regulatory gene or subsequences thereof.

As used herein, "hybridization complex" includes reference to a duplex nucleic acid sequence formed by selective hybridization of two single-stranded nucleic acids with each other.

As used herein, "amplified" includes reference to an increase in the molarity of a specified sequence. Amplification methods include the polymerase chain reaction (PCR), the ligase chain reaction (LCR), the transcription-based amplification system (TAS), the self-sustained sequence replication system (SSR). A wide variety of cloning methods, host cells, and *in vitro* amplification methodologies are well-known to persons of skill.

As used herein, "nucleic acid sample" includes reference to a specimen suspected of comprising RG resistance genes. Such specimens are generally derived, directly or indirectly, from lettuce tissue.

The term "antibody" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments or synthetic or recombinant analogues thereof which specifically bind and recognize analytes and antigens, such as a genus or subgenus of polypeptides of the invention, as described *supra*.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims.

5

EXAMPLES

The following examples are offered to illustrate, but not to limit the claimed invention.

Example 1

10 Example 1 describes the use of PCR to amplify RG genes from lettuce.

Multiple primers with low degeneracy, particularly at the 3' end, were designed based on the sequences of two known resistance genes from tobacco and flax.

DNA Templates

15 Lettuce genomic DNA was extracted from cultivar Diana and a mutant line derived from cultivar Diana using a standard CTAB protocol. To generate cDNA templates, RNA was isolated from cultivar Diana and the mutant following standard procedures; first strand cDNA was synthesized using Superscript reverse transcriptase from 1 Φ g total RNA as specified by the manufacturer (Life Technologies). BAC (bacterial artificial chromosome) clones from the *Dm3* region were isolated from a BAC library of 20 over 53,000 clones using marker AC15 that was known to be closely linked to *Dm3*. Bacterial plasmids containing clones of *L6* and *RPS2* were used as positive controls.

PCR with degenerate oligonucleotide primers

25 Oligonucleotide primers were designed based on conserved motifs in the nucleotide binding sites (NBS) of *L6*, *RPS2*, and *N*. Eight primers were made corresponding to the GVGKTT motif in the sense direction; each had 64-fold degeneracy. Six primers were made to the GPLAL motif in the anti-sense direction; with either 16 or 256-fold degeneracy (Table 1).

30 Oligonucleotides included 14-mer adaptors of (CUA)₄ at the 5' end of the sense primers and (CAU)₄ at the 5' end of the antisense primers to allow rapid cloning of the PCR products into pAMP1 (Life Technologies).

PCR amplification was performed in 50 μ l reaction volume with 1 μ M of each of a pair of sense and antisense primers. The templates were denatured by heating to 94EC for 2 min. This was followed by 35 cycles of 30 sec at 94EC, 1 min at 50EC, 2 min at 72EC, with a single final extension of 5 min at 72EC. 25 ng of genomic DNA or cDNA was used. BAC clones as templates required less. The final dNTP concentration was 0.2 mM; MgCl₂ was 1.5 mM.

- 5 Forty-eight combinations of sense and antisense primers were tested on a panel of nine templates consisting of two genomic DNA samples, two cDNA preparations, three BAC clones and plasmids containing *L6* and *RPS2* as positive controls.
- 10 Amplification from *L6* and *RPS2* resulted in fragments of 516 and 513 respectively. Seven combinations of primers resulted in fragments of approximately this size with multiple templates (Table 2). Primers that gave RLG products were: PLOOPAA, PLOOPAG, PLOOPGA, PLOOPGG, PLOOPAC, GLPL3, GLPL4.

15

20 (Intentionally left blank)

25

Table 1

DEGENERATE PRIMER SEQUENCES for NBS PCR

Sense primers based on GVGKTT amino acid sequence from L6, N and rps2 PLOOP motif:

PLOOPAG 5' GGN GTN GGN AAA ACG AC 3'

PLOOPAA 5' GGN GTN GGN AAA ACA AC 3'

PLOOPAT 5' GGN GTN GGN AAA ACT AC 3'

PLOOPAC 5' GGN GTN GGN AAA ACC AC 3'

PLOOPGG 5' GGN GTN GGN AAG ACG AC 3'

PLOOPGA 5' GGN GTN GGN AAG ACA AC 3'

PLOOPGT 5' GGN GTN GGN AAG ACT AC 3'

PLOOPGC 5' GGN GTN GGN AAG ACC AC 3'

Antisense primers based on GPLPLAL amino acid sequence:

GPLL1 5' AGN GCN AGN GGN AGG CC 3'

GPLL2 5' AGN GCN AGN GGN AGA CC 3'

GPLL3 5' AGN GCN AGN GGN AGT CC 3'

GPLL4 5' AGN GCN AGN GGN AGC CC 3'

GPLL5 5' AAN GCC AAN GGC AAA CC 3'

GPLL6 5' AAN GCC AAN GGC AAT CC 3'

TABLE 2. Characteristics of RLGs isolated from lettuce.

	Template	Primers	Number ^a	Size ^b (bp)	Copy number ^c	Dm linkage
5	RLG1 genomic DNA	PLOOPGA+GLPL6	6/6	522		DM4, DM13
	cDNA	PLOOPGA+GLPL6	1/5			
	genomic DNA	PLOOPAA+GLPL6	5/5			
	cDNA	PLOOPAA+GLPL6	1/1			
10	RLG2 BACH8	PLOOPGG+GLPL3	3/3	510		DM1, Dm3
	RLG3 gemonic DNA	PLOOPGA+GLPL4	3/6	461		Dm5 Dm8
	RLG4 genomic DNA	PLOOPGA+GLPL4	1/6	524		

^a Number of RLG sequences out of total number of clones sequenced.

^b Size of fragment amplified from the nucleotide bindind domain.

^c Estimated copy number from genomic Southern blot analysis and numbers of clones in the BAC library.

Example 2

Example 2 describes the genetic analysis used to obtain a preliminary indication of the linkage relationships of the amplified products and known clusters of resistance genes.

Bulked segregant analysis was performed to obtain a preliminary indication of the linkage relationships of the amplified products and known clusters of resistance genes. DNA from individuals were pooled for each susceptible and resistant bulk. Amplified products were then mapped by RFLP analysis from our intraspecific mapping population. Resistances from four clusters of resistance genes as well as over six hundred markers have now been mapped on this population. Linkage analysis was done using JIONMAP or MAPMAKER mapping programs. Due to a suppression of recombination in the *Dm3* region, sequences were mapped relative to *Dm3* using a panel of deletion mutants that provided greater genetic resolution than the mapping population (Anderson *et al.* 1996). All blots were washed twice at 63EC in 2x SSC/1% SDS for 20 min, followed by one wash at 63EC in 1x SSC/0.1% SDS for 10 or 30 min.

Most of the RLG sequences were analyzed by bulked segregant analysis (BSA) using pools of resistant and susceptible individuals for each of the four clusters of resistance genes. In genomic Southern analyses, all the RLGs revealed numerous fragments of varying intensity. The numbers of bands was highly dependent of the 5 stringency of hybridization. BSA demonstrated that RLG1 was linked to the *Dm4,7* and *Dm13* clusters. Segregation analysis confirmed this linkage.

RLG2 was derived from BAC H8 that was known to be from the *Dm3* region. BSA with RLG2 demonstrated that the polymorphic bands that distinguished the parents of our mapping population mapped to the *Dm1,Dm3* cluster. Several bands 10 absolutely cosegregated with *Dm1* or *Dm3*. To provide finer genetic resolution, RLG2 was also mapped using a panel of *Dm3* deletion mutants. A number of fragments were missing in largest deletion mutant demonstrating that several RLG2 family members are physically located very close to *Dm3*. No fragment was missing in all deletion mutants; however, this is not unexpected as there is extensive duplication within the region.

15

Example 3

Example 3 describes the screening of a bacterial artificial chromosome library.

Over 53,000 BAC clones containing lettuce genomic DNA were screened 20 with two of the amplified products. High density filters each containing 1536 clones were hybridized to ³²P labelled probes. Filters were washed at 65EC with 40 mM Na₂PO₄/0.1% SDS for 5 min followed by 20 min in the same solution.

To isolate additional RLG sequences we screened our genomic BAC library. Clones were identified that hybridized to RLG1 and RLG2. Nearly all the clones that 25 hybridized to RLG2 also hybridized to marker AC15 that had already been shown by deletion mutant analysis to be clustered around *Dm3*. This provided further evidence for clustering of RLG2 sequences.

Using primers conserved within each family, part of the NBS was amplified from each unique BAC clone and sequenced. This revealed that members within each 30 family varied from 64% identical at the deduced amino acid level. The most divergent members only weakly cross-hybridized to each other. Currently, RLG sequences are

considered to be part of the same family of sequences if they are at least 55% identical at the deduced amino acid level and map to the same region of the chromosome.

Example 4:

5 Example 4 describes the cloning, identification, sequencing and characterization of RG polynucleotide sequences; including use of RG sequences from plasmid and PCR products.

Doubled stranded plasmid DNA clones and PCR products were sequenced using an ABI377 automated sequencer and fluorescently labelled di-deoxy terminators.

10 Sequences were assembled using Sequencher (Genecodes), DNASTar (DNASTar) and Genetics Computer Group (GCG, Madison, WI) software. Database searches were performed using BLASTX and FASTA (GCG) algorithms.

15 Sequences flanking the NBS region for RLG2 and for some of RLG1 were obtained by a series of IPCR and the products sequenced directly. IPCR worked less well for RLG1. Therefore RLG1 was subcloned from a BAC clone into pBSK (Stratagene) and the double stranded plasmid sequenced by long range sequencing.

Initially, a total of 30 clones were sequenced. Three of these seven primer combinations yielded sequences that comprised continuous open reading frames with sequence identity to the NBS of known resistance genes. Seven out of 10 clones amplified from genomic DNA with the primer pair PLOOPGA/GLP6 were 522 bp long; they were identical to each other and named RLG1. All six clones amplified from genomic DNA or cDNA using the primers PLOOPAA/GLP6 were similar/the same as RLG1. All three clones sequenced from BAC clone H8 were 510 bp long, identical to each other but different from RLG1 and were therefore designated RLG2. The 11 clones sequenced from four other primer combinations had no similarity to any NBS motifs and therefore were not studied further. Therefore, sequencing resulted in the identification of clones containing NBS motifs representing four RLG sequences.

30 Comparison of the deduced amino acid sequences of RLG1 and RLG2 to those of known resistance genes revealed that RLG1 and RLG2 are as similar to each other as they are to resistance genes from other species and that this is the same level of identity shown between the known resistance genes (Table 3). The percent identity (upper quadrant) and percent identity (lower quadrant) were determined using the MEGALIGN

routine of the DNASTAR package. Identity refers to the proportion of identical amino acids; similarity refers to the proportion of identical and similar amino acids and takes into account substitutions of amino acids with similar chemical characteristics. RG1 and RG2 are as similar to each other and to cloned resistance genes as cloned resistance genes from 5 a variety of species are to each other. L6, resistance to *Melampsora lini* in flax (Lawrence et al., 1995). N, resistance to tobacco mosaic virus in tobacco (Whitham et al., 1994). PRF, required for resistance to *Pseudomonas syringae* in tomato. RPS2, resistance to *Pseudomonas syringae* in *Arabidopsis thaliana* (Bent et al., 1994; Mindrinos et al., 1994). RPM1, resistance to *Pseudomonas syringae* pv. *maculicola* in *A. thaliana* (Grant et al., 10 1995). The initial RG1 and RG2, sequences were amplified from lettuce using degenerate primers.

15 **Table 3**

IDENTITIES OF

RESISTANCE GENE HOMOLOGUES

	RG1	RG2	RG3	RG4	N gene	RPS2
Lettuce	RG1	***	22.7	15.0	29.2	25.4
Lettuce	RG2		***	32.2	21.6	22.7
Lettuce	RG3			***	17.2	15.0
Lettuce	RG4				***	32.8
Tobacco	N gene					44.3
Arabidopsis	RPS2					22.7
					***	21.6

25

The regions homologous to the primers are included in this analysis as the genomic sequences for RLG1 and RLG2 were determined by IPCR. Interestingly, the genomic sequences for RLG1 exactly matched that of the primers used.

To obtain further evidence that we had amplified resistance genes, we 30 amplified the regions flanking the NBSs of RLG1a and RLG2a by IPCR of BAC clones. These products were then directly sequenced without cloning to minimize the introduction of PCR artifacts. Sequence analysis of the 5' regions failed to detect any homology to known resistance genes. However, the sequence of the 3' region contained leucine-rich

repeats (LRRs). When this sequence was used to search GENBANK using BLASTX, it detected identity to the *Arabidopsis* resistance gene, *RPS2*. This region does not contain as regular LRRs as in some resistance genes; however, the repeat structure seems to be consistent with that of the flax resistance gene, *L6*. Therefore, the presence of an LRR 5 region is further evidence that the sequences we amplified using degenerate oligonucleotide primers are probably resistance genes.

The sequences of the IPCR products also provided the genomic sequences of the regions complementary to the sequences of the degenerate oligonucleotide primers. The genomic sequences for RLG1 were identical to one of the primers in the mixture. 10 The RLG sequences are resistance genes as supported by three criteria: the presence of multiple sequence motifs characteristic of resistance genes, genetic cosegregation with known resistance genes, and their existence as clustered multi-gene families. The presence of LRR regions in a similar position relative to the NBS as in cloned resistance genes provides stronger evidence than relying solely sequence similarity between NBS regions. 15 The clustering of RLG sequences at the same position as the known clusters of resistance genes make them strong candidates for encoding resistance genes. The hybridization patterns and genetic distribution of the RLG sequences are similar to that of cloned resistance genes in other species. Most of these hybridize to small multigene families and preliminary genetic evidence indicates that they are clustered in the genome. Therefore, 20 the degenerate primers that we designed from other resistance genes seemed to have been specific enough to amplify resistance genes rather than P-loop containing proteins in general.

SEQID NO:1

FIGIA
[Strand]

1 ATCGTAAACGGTTCTACGAG ANGGCTGTCCCTCTTCATC TTTGTCAATATGTCATATTTC TCATNNATTTGCCACATNT
 81 AATTGGGTATTTAAA TTAAATTTCATTCACATGT CATTATGAGTTTCTAT TTATGATTTCACTAAT
 161 ATTAATGTAATAACAATA ATGCATATTATTTCTT TAAATAACGCATATAATAT ATAGATTAATTCATATAAT
 241 ACATAGTTAAACTCATATA ATACATATGTCATCCCCAG TTATTTATGTCATCC TTATGCAAATTTTATTTAT
 321 TTATTAGAGTAGATGATCTT TGTGATTTAAATTAAATTTAAT TGTGCAAATTTAAATTAAATTA TTAATAATCCCACATTGA
 401 ATAAATTAAATTTAAATGGN CCCACCATAGTCATCACT TTTTCAGTCATCAATATCG TGAGTATTCCTTCGTTTC
 481 CACCCCTTAATCAATTTCCA CGGAATGACAGACTCTTACG CGGTTTCTGAATTTCGGTC CGACACTGTCATIGAGGA
 561 GATAATAATCAATGGAGC TGCTCCATGTCATGCTG ATGAAAGGTAATTGATGT GAAGANAATGTCAGCGATCN
 641 ATCTCCATCCGGAACCCACC ACATTATGAGTGTACCA AACCACCTAAACGGYGGAA GTAGRRAKACWRKAAAGTC
 721 TGAAGAATAGATATTITG TCTCATGGGTGACTGGAG AGGGGGTTAGTCATCATT TTCTTGTGANCAGAAAGAATTAA
 801 TCGGTCATCGAATTTTAC ATGCACAAAGGAGTTTAC TCGCAATGTTGTAACAA ATTATTAATCTTATCTT
 881 TTGTTGAAACTCTCAATT GCAACTTCGAACCTTCAACT TTGGGCCACAAATTG TGCGGGCTTAATTAAATCCA
 961 CATATTCACTGTAACAATA ATTCAAATCGATCTGTC ATCCAATTCAACATCTC TTGATAATTGAAATCATTCA
 1041 CGCTTCATCCATTTCATCCA CATCTATACATATACTCTG CTCTTATCATATTAAACGAT GGCTGAAATCGTTCTTC
 1121 CCTCTTGACAGTGGTGTGTT GAAAAGCTGGCATYTGAAGC TTGAGAAGAGTGTGCGCT CAAAAGAATTGAATCTGAG
 1201 CTTAGAAATGAGGAGAC ATTAGACCAATTCGAAGATC TGCTTAACGATGCTTCGG AAGGAAGTAATTAATGAGC
 1281 CGTTAAAGATGGCTGAATG ATCTCCAATTTGGCTTAT GACATAGACGACCTACTTGA TGATTTGCAACTGAGCTG
 1361 TTCACTGAGTTGACCGAG GAGGGTGGAGCCTCCCTCCAG TATGGTAAGAAAACAAATCC CAAGTGTGACAGTTTC
 1441 TCACTAAGTAAATAGGATGCA TGCCAAAGTTAGATGATTTG CCACCGAGTTACAAGAACG TGAGAGGCAAAAATAATCT
 1521 TGGTTAAGTGTGATAACAT ATGAAAAGCCAAAAAATGAA AGGTATGAGGCTCTTGTG AGATGAAAGGGTACTGTG
 1601 GACGTGAAGTGAATAAGAAA AAATTGCTGGAGAAGCTGTTT GGGGATAAAAGTGAATCAG GGACTAAAATCTCAGCATC
 1681 GTGCCATAGTTGGTATGGG TGGAGTTGTTAAACAATC TAGCTAGACTTTGTATGAT GAAAGAAAAGTGAAGGATCA
 1761 CTTCGAACTCAGGGCTTGGG TTGTTGTTCTGATGAGTTG AGTATTCCTAAATCAAGCAG AGTTATTTATCAATCTGTGA
 1841 CTGGGAAAAGAGGAGGAGTGA GAAGACTTAATCTGCTTC AGAAAGCTTAAAGAGAAC TTAGGAACCCAGCTTTCTA
 1921 ATAGTTGGATGATGTTG TGCTGAAAGCTATGGTGTGTTTGGGATAAATAGTGGGGCCA TTCTCTGGGGCTCCCTGG
 2001 AAGTAGAATAATCATGACAA CTGGAAAGGAGCAATTGCTC AGAAAGCTGGCTTTCTCA TCAAGACCCCTCTGGAGGGTC
 2081 TATCACAGATGATGCTTGT TCTTGTGTTGCTCAACACCC ATTGGTGTACCAAATCTG ATTACACATCCAACACTAAGG
 2161 CCACATGGAGAACATGTTGTTG RAAGCAGGAGGACAATGGGA AGAAGATTTGATGGCTTAC CCAAGGTTATTAAAGGACAAA
 2241 AACAGACGGGACAATGGG AGGAGCTGTTGGATAGTGTG ATATGGGGTTAGGAAGAG CGATGAGATTGTCGGCTC
 2321 TTAGACTAAGCTCATGATGAT CTTCCTGCCCTTTGAAGCT RTTTRTTGCAATATGCTCT TGTTTCCCAGGACTATGAG
 2401 TTGACAGGGAGGAGTTGAT TCTATTGTGGATGGCAGAAG GGTTTTGCACCAACACT AYAAACAAAGTCAAAGCAACG
 2481 KTGGGTCTTGAATTTTGT TGACGTTTCAACATGCTCTAA TRRCAAATCSTTGTGTAAGGAAAGGAAAG
 2561 TGCATGACCTAATGATGAT TTGGCTACATTGTTGCTGG AGAATTCTTCTCAAGGTTAG ACATAGAGATGAGAAGGAA
 2641 TTAGGATGSAATCTTGTGA RAAGCAGGAGGACAATGGGA AGTATGTTGGAGATTACATA GGTTACAAAARGTTGGAGCC
 2721 ATTAGAGGGAGCTAAAATT TGAGAACATTTTAGCATTTG TCTGTTGGGGTGGTAGAAGA TTGGAAGATTTTACTTAT
 2801 CAAACAGGTCTGAATGAC WTACTTCAGGATTTACCAATT GTTAACGGCTCTRAKTTGAT TTRRTCTTAYAATAASYRAG
 2881 GTACCARAAATCTGSGGTAG TATGAAACATGCTGGTATC TTAAATCTATCWGRAACTWA ATCACMCATTIACCGGAAWA
 2961 TKTCGAAATCTTATAATT TACARACCTGATTTGTCCT GTGCTGTGAMATTAGTTAA KTTGGCCCAARACCTTCTCAA
 3041 ASCTTAAAATTTCASCAT TTGACATGAGGGTACTCC KAATTTAARAACATGCCCT TARGATTGGTGTGATMIGAAA
 3121 ARTCTCCTAAACTCTTCTTGG TAACATTGGCATAGCAATAA CCGAGCTTAAGAACATTGTCAM AAYCTCCATGGGAAARTTTG
 3201 TATTGGGGCTGGGAAAAA TGGAAAATGCMGTTGGATGC ACGTTAACGGAACCTGCTC AAAAAGGTTWAATGARTTA
 3281 NAAACTGGRWTTGGGGTGAG TRAATTAAATGTTTCCGAA ATGGGACACTTAAAGA AGTCTCTAATGAGTGTG
 3361 CTCATATGGTACTCTANAA AAAACCCANAAATTGCTCA TAGGGGGTATAGAGTTTCCA ATTGGGTTGGTINCACTAA
 3441 GGGTTCTGAAACTAGAGAT GTGTTCTGTTGATGAAAA AGANTGTTTACGTAGTTTC ATCAATCACAAGTGGGAAA
 3521 TAGATGATATTTCAGGGCY TACTGATGAGATGTTGGAGAG GTATGATGAGGTCTCTGG GCGGTAGAAGAAATAAGCAT
 3601 CCATTCTGTAATGAAATAA GATATYTGIGGGAAATCAGAA CGAGGAGGAACCTGTTCT TATGAATTTAAAGAAGTTGG
 3681 ATTTAGGTGAATGTGAAAAT TTGGTGTAGTTAGGGAGAA AAAGGAGGATAATCTATAA TTATAGTGGGAGCAGGCTA
 3761 ACATCTTCTAGGAGGTTGAA TGATGAGGATGAAACAGCT TGGAGCATTCAGGTTTCCA GATAGCATGGGAAATTGTA
 3841 TATGCCACATGTTGATTCAA TNACATCCGCTCTTCCCA ACAGGAGGAGGACAGAAGAT CAAGTCACCTACCATCACTG
 3921 ATTGCAAGAACCTTCTGGAA GAGGAGTTGGGAGGAGGAGA GAGGACAAAGGTGCTTATAA ACTCAAAAATGCAAGATGCTT
 4001 GAATCTAGATATACGTAAT TTGGCCAAATCTGAAATCTA TCAGTGAATTGAGTTGCTTC ATTACACCTGAACGATTATA
 4081 TTATCTAATGTCGGAGTP TGGAGTCATTCTGACCAT GAGTTGCCAAATCTCACCTC CTAAACAGATGAGGGAGAG
 4161 GACAGCGATTTCGACGAA CGGTTACGATTGACTGCC GTCGTTT

SEQ ID NO: 2
 RIGB
 [Strand]

1	AACCGTTCGT	ACGAGAAATCG	CTGTCCCTCTC	CTTCCCTGTA	TATAATGATA	AGAAAAAATA	TGATTAAAGG
71	TTTAAATCCA	AAATCCATT	TTCCACCGG	GATATGATGC	ACTAGCTGTA	GTATGCAAAA	ACAGTATTTAT
141	AAATGCTAAC	CAAACAGCA	GCTAAAGAAC	AAATAATAA	ATGGTTTGA	TGGTCCTTTC	TCCGTACAQT
211	CATTTCTTCC	AAATCCCTAT	CATTCAAC	TACAAGTCT	CCCATTTAG	TTTTTCACTA	TAAGCAATGG
281	CTGAAATCCCT	TGGTTCTGCG	TTCCTTGCGG	TGTCTTGTGA	AAAGCTTGTCT	TCTGAAGGCCT	TGAAGAGGGT
351	TGCTTGCTCC	AAAGTAATTG	ACAAGGAGCT	CGAGAAATTG	AATAGCTCAT	GAATCATAAT	AAAAGCTCTG
421	CTCAATGATG	CTTCTCAGAA	GGAAATAAGT	AAGGAAGCTG	TAAAGAAATG	GTGAAATGCT	CTTCAACATT
491	TGCCCTTACGA	CATAGATGAT	CTACTTGGCG	ATTTCGCAAC	CAAAGCTATC	CATCGTAAGT	TCTCTGAGGA
561	ATACGGGGCC	ACCATCAACA	AGGTACGAAA	TTAAATTCCA	TCTTGTCTCT	CTAGTTTGTG	AAGTACTAAG
631	ATGGCCAAACA	AGATACATPA	TATTACCAAG	AACTTACAAG	AACATTTAGA	AGAGAGAAAT	AATCTTGGAT
701	TATGTGAAAT	TGGTGAAGC	CGAAAACCTC	GAAATAGAAA	ATCAGAGACC	TCTIVGCTAG	ATCCATCTAG
771	TATTCCTGGA	CGCACAGATG	ATAAGGAAGC	GTGCTTCTC	AAGCTATAAG	AACCATTGTA	TAGAAACTTT
841	AGCRATCTTC	CNATAGTGG	TATGGGTGGG	TTAGATAAGA	CCACATTAGG	TAGACTTTTG	TATGATNAAA
911	TGCAAGTGA	GGATCACTTC	GAACCTCAAG	CGTGGGTTTG	TGTCTCTGAT	GAGTTTGATA	TCTTCGGTAT
981	AAGCRAAACC	ATTTTCGAAT	CGATAGAGGG	GGGAACCAA	GAGTTAAGG	ATTAAATCT	GCTTCAGGTG
1051	GCTTTRAAGG	AGAAAATCTC	AAAGAAACGA	TTTCITGTTG	TTCTTGATGA	TGTATGGAGC	GAGAGCTATA
1121	CTGATTGGGA	AATTCTAGAA	CGTCCCATTC	TAGCAGGAGC	ACCAGGAAGT	AAAGTAATCA	TCACAACCCG
1191	CAAGTTGTCG	TTGCTAAACC	AATTGGGTCA	TGATCAACCA	TACCAATTGT	CTGATTGTC	ACATGACAAT
1261	GCTCTCATCCT	TATTTCGICA	ACACGGATT	GGTGTAATAA	GCTTTGATTTC	ACATCCGATA	CTTAAACCAC
1331	ATGGTGAAGG	TATTGTTGAA	AAATGTGATG	TTTGGCATT	GGCTTTGATT	GCACTTGGGA	GGTTATTGAG
1401	GACAAAGA	GATGAGGAAG	AAIGGAAGGA	ACTATTGAA	AGIGAGATAT	GGAGGTTAGG	AAAGAGAGAT
1471	GAGATTATTC	CGGYTCTTAG	ACTAAGCTAT	AATGATCTTT	CTGCTCTTTC	GAAGCAGTTG	TTTCCATATT
1541	GCTCTCTGTT	CCCCAAAGAC	TAITGTTCA	ACAAGGAGAA	GTGATTTTA	TTATGGATGG	CAGAAGGGTT
1611	TTTGCACAAAT	AAAATAACAA	ACAAGTCAT	CGAACCTTAA	GTCTTTGAT	ATTTCGACGA	CTTGTGTC
1681	AGGTCTTTT	TTCAACATGC	ACTCGATGAC	AAATCGTTGT	TGTTGGTGCA	CGACCTCATG	AATGACTTGG
1751	CCACATCTGT	TGCTGGAGAT	TATTTTTAA	GATTAGACAT	TGAATGAA	AAGGAAGCTT	TGAAAAAATA
1821	CCGACATATG	TCATTGTTT	GTGAGAGTTA	CATGGTTAC	AAAAGGTTCG	AACCATTAA	AGGAGCTAAA
1891	AAATTGAGAA	CTTCTTACG	AATGCTGTT	GGGATGATAA	AAAGTTGGAC	AACATTTTAC	TTATCAAATA
1961	AGGTCTTGA	TGACTTACTT	CACGAATTAC	CATTGGTGTG	AGTTCTAAGT	TTGAGTTATC	TTAGCATCAA
2031	GGAGGTCACCT	GAATAATAG	GCAATTGAA	ACACTTGGCG	TATCTTAATT	TATCACACAC	GAGTATCACA
2101	CATTTCACAG	AAATGTCCTG	CAATCTTAC	AACTTACAAA	CATTGATCC	TGTTGGCTGT	TGTTTATAAA
2171	CCAAGTTTCC	CAACAATTTC	TTAAAGCTTA	GAATTTCAG	GCATTGGAC	ATTAGCGATA	CTCCCGTTT
2241	GAAGAAGATG	TCCTCGGGGA	TGGTGAATT	GAAGAACCTA	CACACYCTCT	CCAAGCTCAT	TATGGAGGT
2311	AAAAATAGAC	TAACGAGCT	TAAGAACCTA	CAAAATCTCC	ATG		

RLG1b - Diana
[Strand]

1 TACTACTACT AGAATTCCGGT GTTCGTAAGA CGACTCTAGC TAGACTTTG TATGAGGAAA TGCAAGGGAA
71 GGATCACTTC GAACTTAAGG CCTGGGTATG TGTTTCTGAT GAGTTTGATA TCTTCAATAT AAGCAAAATT
141 ATCTTACAAT CGATAGGTGG TGGAAACCAA GAATTACGG ACTTAAACCT CCTTCGAGTA GCTTTAAAG
211 AGAAGATCTC AAAGAAAAAGa TTTCCTCTTG TTCTTGATGA TGTTTGGAGT GAAAGCTATA CCGATTTGGGA
281 AATTTAGAA CGCCCATTTC TTGCAGGGC ACCTGGAAGT AAGATTATTA TCACCACCCG GAAGCTGTCA
351 TTGTTAACCA AACTCGGTAA CAATCAACCT TACACCTTT CGGTTTGTC ACATGAGAAT GCTTTGTCTT
421 TATTCTGTCA GCATGCATTG CGTGAAGATA ACTTCATTTC ACATCCAACA CTAAACCCAC ATGGCGDAGG
491 TATTGTTGAA AAATGTGATG GttGCCATT GGCAATTGTGCG ACATGATGAT GATG

SEQ ID 137

SEQ ID NO:3

RLGIC

[Strand]

1 TCCCCGTGCAA CGTNTATCAT TCAGAAAGNC CCAAAGACCA NAGATNTGTT TAANGNTGNT TNTCAGAAGG
71 AAGTAATTGA TGAAGCTGTN AAAAGATGGC TGATTGATNT CCAACAATTG CCTTACCGACA CTGANGACNA
141 ACTTGATGAT NTCCCAACAG AAGCTATICA TCGTGAGTTG ATCCGTGAAA CTGGAGCTTC CNCCAGCAJC
211 GTAAAGRAAGC TARTCCCAG TTGTTGCCACA AGTTTCTCAC AAAGTAATAG GATGCATGCC AGGTTAGATG
281 ATAATTGCCGC TAAGTNACAA GAACCTGGTAG AGGCAGAAAAA TAATCTTGGT TTAAGTGTGA TAACATACGA
351 AAAACCCAAA ATIGAAAGAG ATGAGGGGTN TTGTTGAGAT GCAAGTGGTA TCATGGACG TGAGATGAT
421 AAGAAAAAAAT TGCTTCAGAA GCTGTTGGGG GATACTTATG AATCAAGTAG TCAAAACTTC AACATCGTGC
491 CCATAGTTGG TATGGGTGGG GTAGGTTAAA CAACTCTACG TAGACTTTTG TATGATGAAA AAAAAGTGAA
561 GGATCACTTC GAACTCAGGG TTGGGTITG TGTTTCTGAT GAGTTTCAGTG TICCCAATAT AACAGAGTT
631 ATCTTATCAAT CTGTGACTGG TGAAACAAA GAATTTCAGG ATTAAATCT GCTTCAAGAA GCCCTTAAAG
701 AGAAAATCTCA GAACAAACTA TTCTAATAG TTGTTAGATGA TGATGGTCT GAAAGCTATG GTGATTGGGA
771 GAAATTAGTG GGCCCATTTTC ATGCTGGGAC TTCTGGAAGT AGAATAATCA TGACTACTCG GAAGGAGCAA
841 TTACTCAAAC AGCTGGGTTT TTCTCATGAA GACCCCTCTGC ATAGTATAGA CTCCCTGCAA CGTCTATCAC
911 AAGAAGATGC TTGTCCTTG TTTCTCAAC ACGCATTTGG TGACCTAAC TTGATTCAC ATCCAACACT
981 AGGCCATAT GGGGAAACAGT TTGTGAAAAA ATGTGGGGGA TTGCTTGG CCTTGT

SEQ ID NO:4

RLGID

[Strand]

1 CNTACCCCTTC TACGGAGATCG CTGTCCCTCC TCGATCTGCT TAACCATGCT TCCCAGAAGG AAGTNACTAA
71 TGAAGCCGTT AAAAGATGCC TGAATGATCT CCAACATTG GCTTATGACA TANACGACCT ACTTGATGAT
141 CTIGCAACAS AAAGCTATTG NTCSTGAGTT GACCGANGAA GGTGGAGCCT CCACCACTAT GGTAAGAAA
211 CTAATCCCAA GTTGTGAC AAGTTTCTCA CAAAGTTATA GGATGCATGC CRAGTTAGAT GATATTGCCA
281 CCAGGTACAA AGAACTGGTA GAGGCAAAA ATAATCTGG TTAAAGTGTG ATAACATATG AAAAGCCCAA
351 AATTGAAAGG TATGAGGCAT CTTTGGTAGA CGAAAGTGGT ATTGTTGGAC GTTNAAGATGA TNAGAAAAAA
421 TTGATGGAGA AGCTGTGGG GGATAAAGAT GAATCCGGAG TCNAAACTTC ACCATCCTGCG CCATAATTGG
491 TATGGGTGGA GTTGGCNAAA CAACTCTAGC TAGACTCTTG TTIGATGAAA AGACAGTGAA GGATCACTTC
561 GAACTCAGGG CTGGGGTTTG TGTTTCTGAT GAATTCAAGTA TTCTCAACAT AACCAAAGTT ATCTATCAAT
631 CTGTGACCGG GGAAAAGAAA GAGTTTGAAG ACTTAAATCT GCTTCAGAA GCTCTTAGAG GGAAACTACAA
701 AAACAAACTA TTCTAAATAG TTGATGGATGA TGATGGTCC GAAAGCTATG GTGATTCGGG GAAATTAGTG
771 GGCCCCTTTC ATGCTGGGAC TTCTGGAAGT AGAATAATCA TGACTACTCG GAAGGAGCAA TTACTCAAAC
841 AGTTGGGTTT TTCTCATCAA GACCCCTCTGC GTTGTATAGA CTCCCTGCAA CGTCTATCAC AAGATGATGC
911 TTIGTCTTTG TTGCTCAAC ACGCATTTGG TGWCCA

RIGHT
[Strand]

1 TCTAGCTAGA CTTTGTATG ACGAGATCCA AGAGAGGAT CACTTCGAAC TCAAGGCGTG GGTTTGTGTT
71 TCTGATGAGT TTGATATATT CAATATAAGC AAAATTATTT TCCAATCGAT AGGAGGTGGA ACCAAGAAAT
141 TTAAGGACTT AAATCTCCCTT CAAGTAGCTG TAAAAGAGAA GATTCAAAG AAACGATTTTC TACTTGGTCT
211 TGATGATGTT TGGACTGAAA GCTATGCCGA TTGGGAAATT CTGGAAACCC CATTCTTGC AGGGGCAGCC
281 GGAAGTAAAAA TTATCATGAC GACCCGGAAG CAGTCATTGC TAACCAAACG CGGTTACAAG CAACCTTACA
351 ACCTTTCCGT TTTGTCACAT GACAGTGCTC TCTCTTATT CTGTCAGCAT GCATTGGGTG AAGATAACTT
421 CGATTCACTT CCAACACTTA AACCACATGG CGAAGGCATT GTTGGAAAAT GTGCT

SEQ ID NO:5

RLG1F
[Strand]

1 ATTTTCNGCT CNTAAACAAAN AAAAGCAATG GCTGAATCT TCTCTTCNGC ATTCTAGACC AGTATTCTTT
71 GAAAAGNTGG CTTCTGAAGC CTTGAAGAAG ATCGCTCGCT TCCATCGGAT TGATTCTGAG CTCAAGAAC
141 TGAAGGGTC ATTAATCCAG ATCAGATCTG TGCTTAATGA TGCTTCTGAG AAGGAATAAA GTGATGAAQC
211 TGTTAAAGAA TGGCTGAATG GTCTCCAACA TTGCTTAC GACATAGACG ACTCTTGA TGATTGGCA
281 ACCGAAACTA TGCAATCGTGA GTTGACCCAC GGATCTGGAG CCTCCACCAG CTGTAAAGAA AGATAATCCC
351 AACTTGTGG ACAGATTCCT CACTAATG TAGAGATCGT AACAGTGTAG ATAATATTAC CATCAAGTTA
421 CAAGAACTGG TAGAGGAAAA AGATAATCTT GGCTTAAAGTG TGAAAGGTGA AAGCCCCAAA CATAACCAACA
491 GAAGGATTACA GACCTCTTAC GTAGATGCAT CTAGCATTAT TGGTCTGTGAA GGTGATAAGG ATGCTATTGCT
561 CCATAAGCTG CTGGAGGATG AACCAAGTGA TAGAAACTTT AGCATCGTGC CAATAGTTGG TATGGGTGGT
631 GTGGGTAAAGA CGACTCTAGC TAGACTTTTG TATGACCGAGA TGCAAGAGAA GGATCACTTC GAACTCAAGG
701 CGTGGGTTTG TGTTCTGAT GAGTTTGA TCTTCAATAT AACCAAAGTT ATCTTCCAAT CGATAGGTGG
771 TGGARACCAA GAATTAAAGG ACTTAAATCT CTTCAAGTA GCTGTTAAAG AGAAGATTTTC AAAGAAACGA
841 TTTCCTYTTG TTCTGGATGA TGTTGGAGT GAAAGCTATA CAAAGATGGGA AATTCTAGCA CGTCCATTTC
911 TTGCAGGGGC ACCAGGAAGT AAGATTATCA TGACGACCCG GAAGTTGTCC TIGCTAACCA AACTCGGTTA
981 CAATCAACCT TACAAACCTT CSGTTTGTCA ACATGATAAT GCTYGTCTT TATTCTGTCA GCAYGCATTG
1051 GGTGRAGATA ACTTCGATTG ACATCCAACA CTTAACACAC ASGGTGAAG TATTTGTGAA AATGTGAGC
1121 GTTTTACCAATT GGCCTTTRATT GCACCTGGGA GRTTGTGAR GACAAAACA GATGAGGAAG AATGGAARGA
1191 AGTGTGAAAT AGTGAAATAT GGGGGTCAGG AAAGGGAGAT GAGATTGTTC CGGCTCTTAA ACTAAGCTAC
1261 AATGATCTCT CTGCCTCTTT GAAGAAGTTG TTGCTAATCT GCTCTTGTG CCCAAAGAC TATGTGTTGG
1331 ATAAGGGAGGA TTGATTTTG TTGTTGGATEC CAGAACGGTT TMGCCACAA TCAACACAA GCAAGTCBAT
1401 GGAACGCTTG GGHCATGAAG GTTTGATGA ATTGTGTCA AGATCATTTT TTCAACATGC CCCTGATGCC
1471 AAATCGATGT TTGTGATGCA TGACCTGATG AATGACTTGG CHACATCTGT TGCTGGAGAT TTTTTTICAA
1541 GGATGGACAT TGAGATGAAG AARGAATTAA GGAAGGAAGC TTGSAAGAAG YAYCGCCATA TGTCATTTGT
1611 TTGTGAKGAT TACATGGTCA ACACAAAGTT CRAGCCATTS ACAAGGAGCT AG

SEQ ID NO: 6

RLG1G
[Strand]

1 GTGA~~A~~GGATC ACTTCGAAC T CAGGGCTTGG GTTTGTGTTT CTGATGAATT TAATATCCTC AATATAAGCA
71 AAGTA~~A~~TTTA TCAATCTGTA ACCGGGGAAA AAAAGGAGTT TGAAGACTTA AATCTGCTTC AAGAAGCTCT
141 TAAAGAAAAA CTTTGGAAATC AGTTATTCT AATAGTCTG GATGATGTGT GGCTGAAAG CTATCGTGAT
211 TGGG~~A~~GAAAT TAGTGGGCC ATTTCCTCG GGGTCTCCCTG GAAGTATGAT TATCATGACA ACTCGGAAGG
281 AGCAATTGCC AAGAAGCTG GGTTTCCCTC ATCAAGACCC TTGCAAGGT CTATCACATG ACGATGCTT
351 GTCTTGTTT GCTCRAACACG CATTGGTGT ACCA

SEQ ID NO:7

RLG1H
[Strand]

1 TCTAGCTAGA CTTTTGTATG AGGAAATCCA AGGGAAAGGAT CACTTCGAAC TCAAGGGTGC GGTATGTGTT
71 TCTGATGAGT TTGATATCTT CAATATAAGC AAAATTATCT TACAATCGAT AGGTGGTGGA ACCAAGAAT
141 TTACGGACTT AAACCTGCTT CAAGTAGCTT TAAAAGAGAA GATCTCAAAG AAAAGATTTTC TCTTGTCT
211 TGATGATGTT TGGAGTGAAA GCTATACCGA TTGGAAATT CTAGAACGCC CATTTCCTTGC AGGGGCACCT
281 GGAAGTAAAGA TTATTATCAC CACCCGGAAG CTGTCATIGT TAAACAAACT CGGTTACAT CAACCTTACA
351 ACTTTGGT TTGTCACAT GAGAATGCTT TGTCCTTATT CTGTCAGCAT GCAATTGGGTG AAGATAACTT
421 CAATTACAT CCAACACTTA AACCAACATGG CGAAGGTATT GTTGAAAAAT GTGAT

SEQ ID NO:8

RLGI
[Strand]

1 TCTA~~S~~TAGA CT~~TG~~TGTATG ATGAGATGCC AGAGAAGGAT CACTTTGAAC TCAAGGGTGT GGTATGTGT
71 TCTGATGAGT TTGATATAATT CAATATAAGC AAAATTATTT TCCCATCGAT AGGAGGTGGA AACCAAGAAAT
141 TTAACGGACTT AAACCTCCCT CAAGTAGCTG TAAAAGAGAA GATTTTAAAG AAACGGATTTC TTCTTGTGT
211 TGACGACGTT TGGAGTGAAA GCTATGCCGA TTGGAAATT NTGGAACGCC CATTTCTTGC AGGGGCAGOC
281 GGA~~A~~TAAAAA TTATCATGAC AACCGGAAAG CAGTCATTGC TAACCAAAT CGGTTACAAG C~~A~~ACCTTACA
351 ACCTTTCCGT TTTCACAT GACAGTGCTC TGTC~~TTT~~ATT CTGTCAGCAT GCAT~~TTGG~~TG AAGGTA~~CTT~~
421 CGATT~~CACAT~~ CCAACACTTA ACCACATGG CGAAGGCATT GTT~~GAAA~~AT GTGCTGGATT GCCAT~~TGG~~CA
491 TTGTCGACAA

SEQ ID NO. 9

RLGLJ
[Strand]

1 TACTACTACT AGAATTCCGGT GTTGGTAAGA CGACTCTAGC TAGACTTTG TATGAGGAAA TCGAAGGGAA
71 GGATCACTTC GAACCTTAAGG CGTGGGTATG TGTTTCTGAT GAGTTTGATA TCTTCATAAT AAGCAAAATT
141 ATCTTACAAT CGATAGGTGG TGGAAACCAA GAATTTACGG ACTTAAACCT GCTTCGAGTA GCTTTAAAAG
211 AGAAGATTC AAAGAAAAGA TTTCCTCTTG TTCTTGTATGA TGTTTGGAGT GAAAGCTATA CCGATTGGGA
281 AATINTAGAA CGCCCATTTC TTGCAGGGGC ACCTGGAACT AAGATTATTA TCACCACCCG QAAGCTGTCA
351 TTGTTAAACA AACTCGGTAA CAATCAACCT TACACCTTT CGGTTCCTTC ACATGAGAAT GCTTTGCTT
421 TATTCCTGTCA GCATGCCATTG CGTGAAGATA ACTTCAATTAC ACATCCAACA CTAAACCCAC ATGGCGDAGG
491 TATTGTTGAA AAATGTGATG GATTGCCATT GGCATTGTGCG ACATGATGAT GATG

SEQ ID NO:10

RLGIA a.a.

IVTVRTR?LSLLHLLSYVIFS?!PH?ILWLF.INFYSTCHFMSFSILLSFT.YLNVITINAYLFFFK.THIYR
LKSNT.VKLI.YICSSPVLYVSSLIYLLFIY.SR.SL.Y.KFNLFKI.NY..SHNLNKIKKNGPTISPSLFQLNIV
SILLRFHPNQYFQRMTDSYGVSEFAFRHCSLKEINQMELLOCSLLMKGELYVK?MSAI?LHPEPTTLSV
YHQTTQNNGGSR?T?KS.RIDYFCPHGLTEERV.FIFL?KNYRSIEFLHRQRSFTSQCFVKQFLIFLSFR.NS
SIATCNLQLLGPQICGGR.FNPHIHCKQ.FKSISVHPIHQHLLIEIIIHASISSTSILYSLLLSD.TMAEIVLS
AFLTUVFEKLA?EALKKIVRSKRIESELKKLKETLDQIQDLLNDASQKEVTNEAVKRWLNDLQHLAYDID
DLLDD?ATEAV?RELTEEGGASSSMVRKLIPSCCTSFSQSNSRMHAKLDDIATRLQELVEAKNNLGLSVI
TYEKPKIERYEASLVDESGTVGREDDKKKLEKLLGDKDESGSQNFISIVGMGGVGKTLARLLYDEK
KVVDHFELRAWVCVSDEFSPNISRVIYQSVTGEKKEFEDLNLLQEARKEKLRNQFLIVLDDWSESY
GDWEKLVGPFLAGSPGSRIIMTRKEQLLRKLGFSHQDPLEGLSQDDALSLFAQHAFGVPNFDSHPTLR
PHGELFVKKCDGLPLALRTLGRLLRTKTDEEQWKEELDSEIWRLGKSDEIVPALRLSYNDLSA?LKLLFA
YCSLFPKDYEFDKEE?LLWMAEGFLHQPT?NKSQRLGLEYF?ELLSRSFFQHAPN?KSLFVMHDLMD
LATFVAGE?FSRLDIEMKKEFRM?SLEKHRHMSFVCE?YIGYK?FEPFRGAKNLRTFLALSVGVVEDWK
MFYLSNKVLND?LQDLPLLRVL?L?L?I?VP??VGSM?HLRYLNLS?T?THLPE??CNLYNLQTLIV
SGC?YLV?LPKTF?LKNL?HFDMR?TP?LKNMPL?IGELK?LQRLF?NIGIAITELKNL?NLHGK?CIGG
LGKMENAVGCTLSELVSKV??.NW??G..I.CFPKWEHLKKKSSMK.CLIMVL?KPP?IMSIGGIEFPN
WVGSLRVSETRDVFVMEYK?CFT.FHQSPSGK.MIFSG?TDEMWRGMIG?LGAVEEISIHCSCNEIRYLWE
SEAEASKVLMNLKLLGECECNLVSGLGEKKEDNHNINSGSSLTSFRRLNWRCNSLEHCRCPDSMENLY
MHMCDS?TSVSFPTGGGQKIKSLTITDCKKLSEEELGGRERTRVLINSKMQMLESVDIRNPWPNLKSISEL
SCFIHLNR?YISNCPS?ESFPDHELPNLTS LTDERRGQRFSYERLRFDWPSF

SEQ ID NO:11

RLGIB a.a.

NRSYENRCPLLPVI...EKI.LKV.IQNPLFHR.YDALAVVCKNSIIINANQNSS.ETI.IMV.IVLSPYTHFFQIPII
HTYKCSHIRFSLAMAEILGSAFFAVFFEKLASEALKRVACSKVIDKELEKLNSS.INIKALLNDASQKEIS
KEAVKEWLNALQHLPYDIDDLGLATKAIHRKFSEYGATINKVRKLIPSCFSSLSSSTKMRNKHINITS
KLQELLEEPNNIGLCEIGESRKLRNRKSETS?LDPSSIVGRTDDKEALLKLYEPCDRNFSILPIVGMGGL
DKTTLGRLLYD?MQVKDHFELKAWCVSDEFDFISKTIFESIEGGNQEFKDLNLLQVALKEKISKKRFL
VVLDVVWSESYTDWEILERPFLAGAPGSKVIITRKLSSLNQLGHDQPYQLSDLSHDNALSIFCQHAFG
VNSFDSHPILKPHGEGIVEKCDGLPLALIALGRLLRTKRDEEEWKELLNSEIWRLGKRDEIP?RLSYND
LSASLQLFAYCSLFPKDVFNKEKLILLWMAEGFLHNENTNKSMERL?LEYFDDLLSRSFFQHALDDKS
LFVVDLMNDLATSVAGDYFLRLDIEMKKEALEKYRHMSFVCESYMVKRFEPFKGAKKLRTFLAMPV
GMIKSWTTFYLSNKVLDDLLHELPLLRVLSSLSYLSIKEVPEIIGNLKHLYLNLSHTSITHLPENVCNLYN
LQTLLCGCCFITKFPNNFLKLRLNRHLDISDTPGLKKMSSGIGELKNLHTLSKLIIGGENRLNELKNLQNL
H

SEQ ID NO:12

RLG 1c a.a.

SRAT?IIQK?PKT?D?F????QKEVIDEAVKRWLID?QQLAYDT?D?LDD?ATEAIHRELJRETGAS?SMVRKLIPSCCTSFSQSNSRMRHLDDIAAK?QELVEAKNNGLSVITYEKPKIERDEA?LVDASGIIGRED
DKKKLLQKLLGDTYESSSQNFNIVPIVGMGGVGKTLARLLYDEKKVKDHFELRVVVCVSDEFSVPNISRVIYQSVTGENKEFADLNLLQEALKEKLQNKLFLIVLDDVWSESYGDWEKLVGPFHAGTSGSRIIMTTR
KEQLLKQLGFSHEDPLHSIDLQRILSQEDALSLFSQHAFGVPNFDSHPTLRPYGEQFVKKCGGLPLAL

SEQ ID NO:13

RLG ID

?T?LRDRCPSSICLMLPRRK?LMKPLKDGMISNIWLMT?TTYLMILQ?KAI??ELT?EGGASTSMVRKLIPSCCTSFSQSRYRMHAKLDDIATRLQELVEAKNNLGLSVITYEKPKIERYEASLVDESGIFGR?DD?KKLMEKLLEDKDESGVKLQHLPPIIGMGGVG?TTLARLLFDEKTVKDHFELRAWCVSDEFSILNISKVIYQSVTGEKKEFEDLNLLQEAIRGKLQNKLFLIVLDDWSESYGDWEKLVGPFHAGTSGSRIIMTRKEQLLKQLGFSHQDPLRCIDSLQRILSQDDALSLFAQHAFG?

SEQ ID NO:14

RLGIE

LARLLYDEMQEKDHFELKAWCVSDEFDIFNISKIIFQSIGGGNQEFKDLNLLQAVKEKISKKRFLVLD
DVWSESYADWEILERPFLAGAAGSKIMTRKQSLLTKGYKQPYNLSVLSHDSALSLFCQHALGEDNF
DSHPTLKPHGEGIVEKCA

SEQ ID No: 15

RLGIF

FSA?NK?KQWLKSFF?HSRPVFFEK?ASEALKIARFHRIDSELKKLKRSLIQIRSVLNDASEKEISDEA
VKEWLNGLQHLSYDIDDLDDLATETMIRELTTLDEPPPACCKDNPTCCTDFSLSSKMRNKLNDITIKL
QELVEEKDNLGLSVKGESPCKHTNRRQLQTSLVDASSIIGREGDKDALLHKILLEDEPSDRNFSIVPIVGMSG
VGKTTLARLLYDEMQUEKDHFELKAWCVSDEFDFNISKVIFQSIGGG?QEFKDLNILLQAVKEKISKKR
FL?VLDDVWSESYTEWEILARPFLAGAPGSKIIMTRKLSLLTKLGYNQPYNLSVLSHDNALSLFCQHA
LGEDNFDSHPTLK?GESIVEKCDGLPLALIALGRLL?TKTDEEEWKEVLNSEIWGSGKGDEIVPALKLS
YNDLSASLKKLFAYCSLFPKDVFKEELILLWMAEGFLHQSTTSKSMERLGHEGFDELLSRSSFFQHAPD
AKSMFVMHDLMDLATSVAGDFFSRMDIEMKKEFRKEAL?K?RHMS?VC?DYMV?KRF?P?TRS.

SEQ ID No.:16

R L G I G

VKDHFELRAWCVSDEFNILNISKVIYQSVTGEKKEFEDLNLLQEALKEKLWNQLFLIVLDDWSESYR
DWEKLVGPFFSGSPGSMIIMTRKEQLPRKLGFPHQDPLQGLSHDDALSLFAQHAFGVP

SEQ ID NO:17

R LG 1 +

LARLLYEEMQGKDHFELKAWVCVSDEFDFNISKIILQSIGGGNQEFTDLNLLQVALKEKISKKRFLVLD
DVWSESYTDWEILERPFLAGAPGSKIITTRKLSLLNKLGYNQPNLSQLSHENALSLFCQHALGEDNFN
SHPTLKPHGEGLIVEKCD

SEQ ID NO.:18

R L G I F

LARLVYDEMQEKDHFELKAWCVSDEFDFNISKIIFQSIGGGNQEFKDLNLLOVAVKEKILKKRFLLVLD
DVWSESYADWEI?ERPFLAGAAGSKIIMTRKQSLLTKGYKQPYNLSVLSHDSALSLFCQHALGEGNF
DSHPTLKHGEGIVEKCAGLPLALST

SEQ ID NO:19

RLG 15

EFGVGKTTLARLLYEEMQGKDHFELKAWCVSDEFDIFNISKIIQSIGGGNQEFTDLNLLRVALKEKISK
KRFLLVLDVWSESYTDWEI?ERPFLAGAPGSKIIITRKLSLLNKLYNQPYNLSVLSHENALSLFCQH
ALGEDNFNSHPTLKPHG?GIVEKCDGLPLALS

SEQ ID NO: 20

SEQ ID NO: 21
RLG 2A

71	TTNACACCAT	AAATCTCNA	CCCGNGGGGA	CAAAAACCTA	AAAATGGTCC	ATAATGCNCA	AATCAGNAAG
141	CCTCCAACCT	TANCCNCCTCA	TTINACCTCA	NCTGATGCNC	NNTCCTCNTA	AAAGTCANAT	CCAAGCTTGC
211	CCACACTCAC	ACAAGCTCTA	AAITGCCACCT	CTTCTCTCTTC	AAAAGCACAC	AAGAACACTT	TCAAGCTCAA
281	AAAATGTGAGG	CCATAATGTT	GAACNAGGGT	TAGGGCACAT	TTAGGGTTTT	GCTCTCTGGA	AATGGTGCT
351	CGCCCAGTGT	ACACTATGGT	CCTTATATAA	CTTCTCACTCC	CACATTAGG	CTTCAACTCT	GAACGTANTA
421	GCGCAACGTA	CTTCCCCITA	ACCCCAACG	TACTCGGTAG	TCTCCCGTC	AANAAATACAC	TCATGAGTAC
491	AACTTGAGGA	AAGAAAAGGA	TCAAGANAT	ACTCAAAAGC	AAACACATTCT	TTTCAGGAC	TAATTTCGAC
561	GGCTAAAAAA	TAAATGGT	TGTGGAAGCC	TCCGGGATGT	TACAATGAAG	TTGANACCTT	TTGAGACCTT
631	AAATGGTGT	ATTTCTTATT	TGTGGCTGAG	CAAGAACACAA	GGTAAAATT	CGTAATCTAC	
701	AATATTAAG	TIGATAAAGT	TCTCTTATT	ATTTACTTG	ATTTACGGGT	AGTTTTT	TCTTACAAA
771	TAATGTTATCAT	GTGTATTAN	ATAGCCACTA	AAATTGACTT	TTTCCAAAAC	ATAATGTCAA	ATGGTGCCTA
841	GATTTTATAT	TGGAAAACAA	ATAATGAATA	TGATGATNCT	GTCTTATTITA	ANCCGAAAAA	ATTATTCATAAT
911	GGATAAAATGT	GTAAATTAA	AGTTGTGATT	TTTNGCATAA	TATAATCAA	TCCNCCTTTG	TNTGGGAGGT
981	ATGAAAAAT	GTTCACCAA	NAACAGGTG	TTINACNTG	AAGGGTIVGG	AAAGGTGAA	AAAAGTAA
1051	TGATATATAG	TAAGCATAAA	ATGTTGATC	OCAGTGAATA	TNATGTTTAA	GGTINATTGT	ATTAATGTT
1121	ATTTTTGATA	TANATTAGGA	ATGAAAATGA	TGTGACTTTAA	ATTTATAAGT	TATNCNACT	GGATTGAAAC
1191	ATTTTTATTA	NAATATAGAA	NCATCCCTT	GCACACCTAA	CATACTTATC	TTTGGTAGTT	TGGTTATTAT
1261	AAAAATGTACC	TIGCTATTAA	ATTTTAAACC	ATTTTAAACC	CATATGTGG	ACGGACTTGA	ATAAATGGGA
1331	AAAAAACTTA	TCTCTTGTG	ATAGGTCAAC	ATGTTGATC	TGTGACTTGC	TATTGAGC	AAACAAAAAA
1401	CTAAAAGAAA	CTATTTGCC	TTTATTAAC	CGGGTAAACC	ATTAAGAAA	ATGTTGTC	ATTAATGCA
1471	GCATGAAAAA	AAATAACTTT	CCATTTTTG	CATCCGGTCA	CAATAATAGA	TTGTCATTTA	TTGTCATTTA
1541	TTTACGGAAA	CTAACCTCTC	TTTTTCTTT	TGGCATCTGA	TCATAAAATA	TTGTTAGTT	TCAGTTAGTT
1611	TTACATTTT	AATACATGTA	AATGTCATA	CCACATGTA	TTCTATPAAA	ACGGAAATGT	ATTTACTTA
1681	TTCCTTGATT	CTTGGCTTC	TTTTTAGTAC	CCAAAACATC	CCCTATCCTA	TCTATTCCA	CTAAAATAAT
1751	AAAAAACTATA	TCCTTCCAT	TGTAGGGATG	TTATAAATT	TGTAAATTGTT	TTTATGCAA	AAAGTGT
1821	TGTTAATCT	GATTAACGAG	ATTCAATT	CAGCATTAA	GGAGAAGTTC	ATCCATCTT	TGGATATGAA
1891	GTGCAAGCCA	AGTTCTTAA	CATGGAAAT	GAGGTCCCTA	TATGTCAAA	AAATAGCAA	TGAGAAATT
1961	TTTAAATTGG	ATCCCCATAA	AGAAAATT	GTTAATGGT	TTTTAATAT	TGGTCATATG	GTCCACGGGA
2031	TGAGCTAAT	ACTAGTTTAT	AAGGGTAA	GGGGGGTTG	TTTATCTTAT	TTATCTTAT	TATTCCTTAT
2101	AGTCAGAATT	AGTAAAAAAA	AATTATAAGA	TAATACCAT	AAGGATAAAA	ATTCATTTA	TTTGGACCAA
2171	AGACCAAAGT	TGTTAGGGG	CIGCTTGT	TTTTTGTGAA	GAGCTGTGCA	ACCACTTTG	TCTGGCCG
2241	ACAGACCAACG	TGCGACATA	TGCCCTCGGA	GAGTGTGTT	TTTTGAAAG	TCCGCGACCC	AAAAAAACGT
2311	CTGGCGGAGG	TCATCTGGC	GCATATATG	GTCACTGTCT	TCAAAGCTT	TCRACCTCA	TTTAAACCA
2381	AAAAAAACAA	GACCACCGGT	TTTTTTTT	TTTTTTTTCT	TTCTCTGTG	CGTCAAAAT	CATTTCATAAT
2451	CTTATGACA	TGAAATTAAAG	TTTGAAAAT	TAATTATTT	CAACAGCTG	AGACGTAAA	AACAAACAGT
2521	CTCTTGTG	CAGACTGTGG	ACATTGGTC	CACCTCTCT	ACCGCAGAGA	CTTGCAGATG	TGGTCCGAG
2591	ACTGCGACAG	TITGGCTTC	AAATAACAA	ACATCACCTA	ATTTGACTAC	ACACACCGGA	CCTCCAAATGT
2661	AAACAAAAAA	AGGTGAAAC	AAAGTTCCT	ATTCTCTTCA	ATCCACGGGC	CATTATGTA	AGAGTTATCT
2731	AAATTTTGT	TGGTAGATC	AGTTCTCAC	TTTAACCGG	GTAAAGTGT	TGIGTGTAG	CGCGCACCTG
2801	AAAGGTGTA	ANGTAACCTC	CAAACGTGAAN	CAANAATCGA	TATGAAGTAT	CAAGTTAGAG	GTTCATAATTG
2871	TGAAGGAACT	AGCTGGAGGT	TGGGGAATCG	AGCTTCCACT	TTAAGGTTAA	AATCCATAAC	CCTAAATGTT
2941	GGTACCCCTCA	TATATCAAAT	TCGGTGT	TTGTAATGAA	AAAACCATGC	TCAAAAAACCC	AGTGTAAAGGC
3011	ACGGTATATG	ACATATTAT	ACTTACTGAT	AAAACATTAT	GATAATTG	GTTTACGT	AGTTAGGATT
3081	CGTACCTTCAA	CCAAATGTAA	TAGTTTTGT	GAGTCATCT	ATGTTATGG	GGATACATAC	TGCAACGGG
3151	ATTTGACTAG	TAATCGAAA	AGTCCTTTA	AATAATT	CTGTTTATAA	TTTATGAA	TTTATGCGA
3221	CATCTAATAT	TAATAGAAT	GTATCTGATA	TTGAAATT	GTCTCTTAA	TGAACATAGA	CCTTTTCCAT
3291	TTACTAATGC	CTAATTATTA	GTTCCTAATC	AATAATT	ATTTCGTT	TTATGCTCT	AGACAAATAA
3361	AAATCCATGA	TITACCTTTA	AATATTAACCA	AAAATGACCA	TTAATAATA	AAAAAAATTCG	ATACCAAAC
3431	CCCCCCCCAT	GCCCAATGTC	TTAATAATTCT	TGATGCTTT	GTCTTCCCT	CTTTCCCT	TTAGTCCTT
3501	ATTCTGGAGA	TGTTGAGAGA	GTTCATAC	AGAAAAATTTC	AAGAGAAAG	CAAAGGTCCA	GTATTCTCT
3571	TTCTCTTCAAT	ATGTTAAAC	TTACAAAGCT	TTTACACG	ATTCACGGT	TTTGTGTAT	TTTTTCICAA
3641	TTGAAACACTAG	ATTCGGACTT	TTGCCCTTGA	TGATTCTCAA	GATATTGCA	GGAGTGTAGA	TGIGTGAAGA
3711	AAAGTGGTGA	ATAGAAAGAG	CAAGTGAATC	CAGATATAGT	ATTCGTAATA	TATGATGATG	AGATAGAGAT
3781	ATGTTAAAC	TGGCTAGAAA	ATITGTTTAA	TTTGAATT	AGGTGTTGA	ATTGAAAGA	TACCAAGCTA
3851	ATTAACATT	AGTATGCTA	AATAGTATTA	AAGAACAA	AACTCTGAT	TTTTTTTCA	TGATTTTCAA
3921	CCTCTTGGTA	CCAAACTAA	TTATAACAAA	ATTGAAATTC	ATTCCTGCA	ATCAATT	ACTTTTGTIA
3991	TTATCCTCAT	GTCTAAATT	GCCACAACTT	TATTTCCTCA	GTCTACATTG	ATTATGAA	GACTATT
4061	ACCAATTACA	TCTTACTTT	ATGGCCAAAG	CTAATACAA	CCGACTAAAC	TAAGGATT	TAGGATGCT

SEQ ID NO: 21
RLG 1A cont.

4131 ATAGTTTGCT CCCCCATTAT AGATTTCTAT CTAATTGTC TATTTGACTA ATTTCAGGTGC CACCACAAAGT
 4201 AAATTCCCTGA AATGGATGTC GTTAATGCCA TTCTTAAACC AGTTGTCGAG ACTCTCATGG TACCCGTTAA
 4271 GAAACACATA GGGTACCTCA TTCTCTGCAG GCAATATATG AGGGAAATGG GTATCAAAT GAGGGGATIG
 4341 AATGCTACAA GACTTGGTGT CGAAGAGCAC GTGAACCGGA ACATAACCAA CCACCTTGAG GTTCCAGCCC
 4411 AAGTCAGGGC TTGGTTTGAA GAAGTAGGAA AGATCAATG AAAAGTGGAA ATTTCCTCA CGGATGTMIG
 4481 CAGTGTGTTTC ATCTTAAAGG TTAGACACCG GGTGGAAAG AGACCTCCA AGATTAATTGA GGACATCCGAC
 4551 ACTGTCATGA GAGAACACTC TATCATCATT TGGAAATGATC ATTCCATTCC TTAGGAAAGA ATTGATTCCA
 4621 CGAAAGCATC CACCTCAATA CCATCAACCG ATCATCATGA TGAGTTCCAG TCAAGAGAGC AAACTTTCAC
 4691 AGAAGCACTA AACGCACTCG ATCTTAACCA CAAATCCAC ATGATAGCC TATGGGAAT GGGGGAGTIG
 4761 CGGAAGACGA CAATGATGCA TCGGCTCAA AAGGTGIGA AGAAAAGAA ATGTTTAAT TTTATAATTG
 4831 AGGGGGTTGT AGGGGAAAAA ACAGACCCCA TTGCTTATCA ATCAGCTGTA GCAGATTAC TAGGTATAGA
 4901 GCTCAATGAA AAAACTAAAC CAGCAAGAAC TGAGAACCTT CGGAAATGGT TTGTCGACAA TTCTGGTGGT
 4971 AAGAAGATCC TAGTCATACT CGACGATGTA TGGCAGTTG TGGATCTGAA TGATATTGGT TTAAGTCCTT
 5041 TACCAATATCA AGGTGTCGAC TTCAAGGTTGT TGTGACATC ACAGAGACAA GATGTTGCA CTGAGATGGG
 5111 AGCTGAAGTT AATTCAACTT TTAATGTCGA AATGTTATAA GAAACAGAA CACAAAGTTT ATTCCACCAA
 5181 TTTCATAGAAA TTTCGGATGA TTGTTGATCCTA GCGTCCTATA ATATAGGAT GAATATTGTA AGGAAGTGTG
 5251 GGGGTCTACC CATGCCATA AAAACCATGG CGTGTACTCT TAGAGGAAA AGCAAGGATG CATGGAAGAA
 5321 TGCACTTCTT CGTTTAGAGC ACTATGACAT TGAAAATATT GTTAATGGAG TTTTAAAT GAGTTACGAC
 5391 AATCTCCAAAG ATGAGGAGAC TAAATCCACC TTTCCTTGT GTGGATGTA TCCCGAARAC TTGATATTIC
 5461 TTACCGAGGA GTTGGTGAGG TATGGATGGG GTTGGAAATT ATTTAAAAAA NTGTATACTA TAGGAGAAGC
 5531 AAGAACCCAGG CTCAACACAT GCATTGAGCG GCTCATTCAT ACAAAATTGT TGATGGAGT TGATGATGTT
 5601 AGGTGCATCA AGATGCATGA TCTTGTTCGT GCTTTGTTTG TGATATGTA TTCTAAGTC GACCATGCTT
 5671 CCATTGTCAA CCATGTAAT ACACCTAGAGT GGCACTCAGA TAATATGCACT GACTCTTGTG AAAGACTTTC
 5741 ATTAACATGC AAGGGTATGT CTAAGTCTCC TACAGACCTG AAGTTCCAA ACCTCTCCAT TTGAAACTT
 5811 ATGCTGAAAG ATATATCATT GAGGTCTCCC AAAACCTTT ATGAAGAAAT GGAGAACCTT GAGGTATAT
 5881 CCTATGATAA AATGAAATAT CCATGCTTC CTCATCACC TCAATGTTCC GTCAACCTTC GCGTGTGTTCA
 5951 TCTACATAAA TGCTCGTTAG TGATGTTGA CTGCTCTGTG ATTGGAAATC TGCGAATCT AGAAGTGTG
 6021 AGCTTGTGTC ATTTCGCTCAT TGACGGTTG CCTTCACCAA TCGGAAAGTT GAAGAACCTA AGGCTACTGG
 6091 ATTGAGCAGA TTGTTATGGT GTTGTGATAG ATTAATGGT CTAAAAAAA TTGGTCAAAC TGGAGGAGCT
 6161 CTATATGACA GTGGTTGATC GAGGTGCAA CGCGATTAGC CTACAGATG AAACACTGCAA GGAGATGGCA
 6231 GAGCGTTCAA AAGATATTAA TGCAATTAGA CTTGAGTTCT TTGAAAACGA TGCTCAACCA AAGAATATGT
 6301 CATTGGAGAA CCTACACCA TTCCAGATCT CAGTGGGGCG CTATTTATAT GGAGATTCCA TAAAGAGTAG
 6371 GCACCTGTTGAT GAAAACACAT TGAAGTTGGT CTCTGAAAAA GGTGAATTAT TGGAAGCTCG AATGAAACGAG
 6441 TTGTTTAAAGA AACACAGGT GTTATGTTTA AGTGTGGAG ATATGATGA TCTTGAAGAT ATTGAGGTTA
 6511 AGTCATCCTC ACAACTTCTT CAATCTTCTT CGTTCAACAA TTAAAGAGTC CTGTCGTTT CAAAGTGTG
 6581 AGAGTTGAAA CACTCTCTCA CACCTGGTGT TGCAAAACACT TTAAAAAAAGC TTGAGCATCT TGAAGTTAC
 6651 AAAATGATA ATATGGAAGA ACTCATACGT AGCGAGGGTA GTAGAAGAAGA GACGATTACA TTCCCCAAGC
 6721 TGAAGTTTTT ATCTTGTGTT GGGCTACCAA AGCTATCGG TTGTCGCGAT AATGTCAAA TAATTGAGCT
 6791 ACCACAACTC ATGGAGTTG AACTTGACCG CATTCCAGGT TTCAACAGCA TATATCCCAT GAAAAGTTT
 6861 GAAACATTTA GTTGTGAA GGAAGAGGTA AATATAAATT TTAAATGCTA ATACATTACA AAGGATCTT
 6931 TCAGTTAAAT CTTTCAAAAT ATATGTTAAT TTGATGTTG GGGTATTAT TGTGGATGG GACTTATAAT
 7001 AAAATGATTAT CTGCAAGGT CTGATTCCTA AGTTAGAGAA ACTCATGTT AGTAGTATGT GGAATCTGAA
 7071 GGAGATATGG CCTTGCCTAAT TTAATATGAG TGAGGAAGTT AAGTTCAAGAG AGATTTAAAGT GAGTAACCTG
 7141 GATAAGCTTG TGAATTGTT TCCGCACAG CCCATATCTC TGTCATCA TCTTGAAGAC TTAAAGTC
 7211 AGAATTGTTG TTCCATTGAA TCGTTATTC ACATCCATTG GGATTGTTG GTGCAACTG GAGATGAA
 7281 CAACACTGAGT GTGTAAGA TTATTAAGT GATCAGTTG GATAAGCTTG TGAAATCTCTT TCCACACAA
 7351 CCCATGCTA TACTGCATCA TCTTGAAGAG CTGAGACTGG AGAATGTTG TCCATTGAA TGTTTATTC
 7421 ACATTGACTT GGATTGTCGT GGTGAAATTG GCGAAGAAGA CAACACATC ACCTTAAGAA ACATCAAAGT
 7491 GGAGTATTAA GGGAGCTAA GANAGGTGTG GAGGATAAAA GTGGAGATA ACTCTCGTCC CCTTGTCTCAT
 7561 GGCTTTCAAT CTGTTGAAAG CATAAGGTT ACNAATGTT AGAAGTTAG AAATGTATTG ACACCTACCA
 7631 CCACAAATT TAATCTGGG GCACCTTGTG AGATTCAAT AGATGACTGC GGAGAAAACCA GGGGAATGAA
 7701 CGAACCGAA GAGAGTGGC ATGAGCAAGA GCAGGTAGG ATTCAATT CACTGTCCTA ATTAAATGAT
 7771 AAGCTCTGC TTTTGAAATA AAAAGGGAC AAACCATTC ATGACTTTAT GTAGCAATAC AAGTCATGTA
 7841 TAAGAGTGTGAC CAACTCTT TTATTTATAA ATGACTACA AAATTTTTT TTTCATTAGA GATCATGTT
 7911 AAATGTGACT AATTCTTCAT CACCTAATT TAGTTGATAA ATCTTTATAA AIGTCACTAG TTACTTTICA
 7981 GTAAAATAAC AATTTATAA AATTATCRAAC AAAAGCATC AACTAAAAAA ATCCCACAAAC CGGTAAATAAT
 8051 TTTAAATAAA AGGATTTAAC ATCTAATACG AACATTCTT TTCTAAACA TGATTTGGAC CAAATATCAC
 8121 CAGCAACTCA AGTTGGAAAT CGATTCACT TAAACTGAA CCAGCATAAT TAGATAGATG AGAGTGTGAG
 8191 CTAAAGTGCC TATATAAGTT CGTTCTCATCT TTTTCTGAA TCTTGTAGC AAGTTGAATG ATTTCCTCT

RLG 2A cont.

8261 TCAAAATGTA TAAAAATCTA CATTATAAAG AGACTAGCTT GAAAAAAAAT GGCTTAGGTG GGCTTGGGT
 8331 TCTCGTAGAT GAAGATCGAA GGGGAGAGTG TGATTTCAAACAC ATCCCTTCATT TTATTTATTT
 8401 ATTATTATTA TTATTTTTG ATATCTTGTCT CATATTGTT ACAGATATGT GAGGTCTATT AATCTTTATA
 8471 AATATATAAA AAAATAATA ACATAATGAA GAAAATTTAA TAAAAGATAA ATTAATAAGG GCACAATAGT
 8541 CTTTTAGGT AAGACAAGGA CCAAACACGC AACAAAAATA AACAGTAGGG ACCATCCGAT TTAAAAAAA
 8611 TAATTAGGGA CCAAAAACAT AAATTCCCCC AAACCATAAG GACCRTTCAT GTAACTTTACT CTTCATTTTC
 8681 GTTTTGTCA TATTGGGTAA ACTATTTTTT TTGTACACAT CTAGGPAACG AACTTGTGA AGTGTCCCA
 8751 TTAGGATGT GACCTACTAC AACCAGTCAT AATAGTCATA TGTGAACACT TCCAACAAT TTATTACTTA
 8821 GGTGTGTACA AAAAACAAAT AGTTACCATG ATGTGAACATG ACTGAAATTTA ATTTACCTT AGCAAGTTAT
 8891 TTTCCTTTT AGGTGTATG GAAACAGTTC CGTGAGACCG TGACTGGAT GGTAGATAAA TTAGTAAAC
 8961 TTAACCCCTTC AATTAACCTA CCTTTTTCTT ATTAACCTAA TTCAACCTA ATATCTGATT CTIGTTTGAA
 9031 AGTAAAGTTGC ATCTTATTTT TTGTATTATC TTGTGCTATA GGATCTCTAG CACTTTTAA TAATTTATTT
 9101 GAAGGTGAAA GATCCAACTA TTATTAATCT GTGGCATT TCCATCATTT GCAACTGTTT CTIGAAAAAA
 9171 AAATACCTAA AATCAAAATA ACCATTTCA AATCCAAAT TATAAGAGAG AATGTAAAT GGACATGGAA
 9241 TCATAATCA TTAACACAGT TCAGTAAACA AGTTGCTAAT TACATTCTT GCTGTGCAGA TTGAAATTCT
 9311 ATCACAGAAA GAGACATTAC AAGAACCCAC TGACAGTATT TCTTAATGTT TATCCCCTC CTGCTCATG
 9381 CACTCTTTC ATAACCTCA GAAACTTATA TTGAACAGAG TTAAAGGAGT GGAGGTGGTG TTIGAGATAG
 9451 AGAGTGGAG AGCAACAACTG AGAGAAATTGG TAACAACTCA CCATAACCAA CAACAACCTA TTATACCTCC
 9521 CAACCTCCAG GAATTGATTC TATGGAATAT GGACAACATG AGTCATGTGT GGAAGTGCAG CAACTGGAAAT
 9591 AAATTC-TCA CTCTTCCAAA ACAACAATCA GAATCCCCAT TCCACAACTC CACAACCTA AAAATTATGT
 9661 ATTCAGAAAAG CATTAAAGTAC TTGTGTTCTG CTCATGGC AGAACCTCTT TCCAACCTAA AGCATATCAA
 9731 GATAAGAGAG TGTGATGGTA TTGGAGAAAGT TTGTTCAACAC AGAGATGATG AGGATGAAGA AATGACTACA
 9801 TTCTACATCTA CCCACACAAAC CACCACTTGT TTCCCTAGTC TTGATTCTCT CACTCTAAGT TTCCCTGGAGA
 9871 ATCTGAGGTG TATTGGTGGA GGTGGTGCCTA AGGATGAAGG GAGCAATGAA ATATCTTICA ATAATACCCAC
 9941 TCGAACTACT GCTGTTCTTG ATCAATTGAA GTTATGCTTT GTACATATTC AATATTAT TTAATTCTT
 10011 TTATTTATTTG CAATATTCTA TAAATATAC AATTATACCACTAAGTAAAT TACCTAGAGG
 10081 GATGGATGCT ATGACACAGC TGCTACACTC CAGAAACTCT AGTAAGGGCA GTTATGGAAAG TTCAATAAAAA
 10151 TGATAATGGC ATCTTTTGAT GGGTAATATA GGCAATTAA GTTTTATTC TTGTTAAAGCA GTATTTAGCA
 10221 AGTACTGGCC AGTAGGAGAG GAGAATATCA CCTTTTGTA AAATCTGGT ATTTGACCCCA GAATTAGTIT
 10291 AAATGTAACA TTATGATAT CAGGGTCTCAT CAGGTGACAG ATATGTAGA ATAGAACAAAT ATATAATPATC
 10361 ACCCCTTACT ATTTTTCTA AGGTATCTT GTAAATATG TGCTTCTTG TTTCATNGA ATNGCATTIC
 10431 GTATATTGTA GGTTTAAAG TGATTTTMC TTCAATAAT CCCGAAATTA ATTAAAAAAA AAAAACAAAA
 10501 AGTACATTTT TGATGTGGAG AGCACTGGTA TCACCTAGTA TATAAAAGC TTGATTTGAA ATTAACCTTC
 10571 TTATACAAA GTTGTGTATA TAGTTTAAAT AGTTTACAT CATTCTTCCA TGTTGTGTG CAGTTGTCTG
 10641 AAGCAGGTGG TTGTTCTTGG AGCTTATGCC AATACGCTAG AGAGATGAGA ATAGAATTCT GCAATGCTT
 10711 GTCAAGTGTAA TTCCATGTT ATGCAGCAGG ACAATGCAA AACGTAAGG AGAGGACAGC GATTCTCGTA
 10781 CGAACGGTTA CGATTCGACT GGCGCGTCTT TTACA

SEQ ID NO: 21

R L G 1 A a.a.

MDVVNAILKPVVTLMVPVKKHIGYLISCRQYMREMGIKMRLGNATLGVEEHVN RNISNQLEVPAQV
RGWFEEVGKINAKVENFPSDVGSCFNKVRHGVGKRASKIIEDIDSVMREHSIIWNDHSIPLGRIDSTK
ASTSIPSTDHHDEFQSREQTTEALNALDPNHKSHMIALWMGGVGKTTMMHRLKKVKEKKMFNFII
EAVVGEKTDPIAQSAADYLGIELNEKT KPART EKL R KWFDNSGGKKILVILDDWQFVDLNDIGLS
PLPNQGVDFKVLLTSRDKDVCTEMGAEVNSTFNVKMLIETEAQSLFHQFIEISDDVDPELHNIGVNIVRK
CGGLPIAIKTMACTLRGKSKDAWKALLRLEHYDIENIVNGVFKMSYDNLQDEETKSTFLLCGMYPE?FD
ILTEELVRYGWGLKF KK?YTIGEARTRLNTCIERLIHTNLLMEVDDVRCIKMHDLVRAFVLD MYSKVEH
ASIVNHSNTLEWHADNMHDSCRKRLSLTCKGMSKFPTDLKFPNLSILKLMHEDISLRFPKNFYEE MEKLE
VISYDKMKYPLL PSSPQCSVNL RVFHLHKCSLVMFDCSCIGNLSNLEVLSFADS AIDL P STIGKLK KLR
LLDLTN CYGVRIDNGVLKKLVKLEELYMTVDRGRKAISLTDDNC KEMAERSKDIYALELEFFENDA QPK
NMSFEK LQRFOQISVG RYLYGDSIKSRHSYENTLK VLEKGELLEAR MNELFKKTEVLCLSVGDMNDLEDIE
VKSSSQQLQSSSFNNLRLV VSKCAELKHFFTPGVANTLK LEHLEVYKCDNMEELRSRGSEEETITFP
KLKFLSLCGLPKL SGLCDNVK II EL PQLM ELEL DDI PGFTSIYPMKKFETFS LLKEEV LIPKLEKLHVSSM
WNLKEIWPC EFNMSEEVKFREIKVSNCDKLVNLFPHPNPM S I LHLEELEV NCGSIESLF NI DLD CAGAIGQ QEDNS ISLR NI
KVENLGKLR?WRIKGGD NSRPLVHG FQS VESIRVTKC?KFRNVFTPTTTNFNLGALLEISIDD CGENR
GND ESEESSHEQE QIEI LS EKETLQEA TDSISNV FPSC LMHSF HNLQKL LNRVKG VEV VFEI ESES P TS
REL VTT HHNQQQPII LPN LQ E LIL WNM DNM SHV WKS CNWNK FFTLPKQQSESPFHNLTTI KIM YCKS IKY
LFSPLMAELLSNLKHIIKIRECDGIGEV VSNR DDEDEEMTTFTSTHTTTLF PSL DLSL TL SFL ENLK CIGGG
GAKDEGSNEISFNNTTATTAVLDQFEVCFVHIQLFI.

SEQ ID NO: 22

RLG 2B

SEQ ID NO: 23

1 AGTTTTTTTT TITCCCAATA TCCATTATA TGCGATTAT TTCTGAAATA ATTTTATCAA AACGCAGGA
 71 ACAATGTAGA ATAATACTGG TATAATTAAAT TATATAAAGT TATTAGGTG AATACCTTGAG CCTACTATAA
 141 TTAAATTATC ATAATTGAA AATCATCAA TTGATTCTCA TGATATTAA TGTTATCAGA TAATTAAATAA
 211 TATGTG-GCC ACACAAATCC ACATCATCAC ACACCCCAAC CTTATGTCGG CTACCTTCACC ACTTGCAATGA
 281 TCCCCACATC TTCCCAACCC CACCGAGGAC TTGGGTCCTC CTTAATATAT CAATTATTIT CTGTAAGTAT
 351 TTATTGTTGT AAATGTTGAA TGTCATTTTA CCTTTTTCT AATATATACA GAAACATAAA TTTAAATGA
 421 AATTCAACTG CGTTTACATT TTGCACTAAA AAAAAGACT GTACTGTTG CAATATTITA CTTATAACCT
 491 GATTAATTAA TTAAAGCTTA ATTCATAAT TTGCACTTAGG TTGTAATTIT GTGTTTTATA GGGAGGGTGA
 561 GGGTCACCGG GAATCAAAGC ACTTATGTA AAGCAGGGGA ATACAAAAAA ATTACTCTGA AACAAATTIT
 631 ATTCAATTAA AGTGAGATAA TAATGTTCTG ATTAGATTAT GAGAAGTAGG AGATTTAAGT GATATATCCC
 701 ATTTAAAAGA AATTGCAATT TTAATTGTTG ATCTCTTGAT GATGACAAAAA TTAACTCGTG ACAGGTTATA
 771 TATCATATAC AAAATGAGTG GCTATGCTT CGCTTCTCAA AAAGCAATT TAGTTATAC ACACCTACAA
 841 ATTTTAAAGG GGGTTAACAA TCTCAAAATA CTTGATAGT AATTATATAA ATATGCATT AACCCTCTAA
 911 AGAAAATGCT ACTAAGCTTG GACCATCTCA GAATTACAAT CATACCTTC CCCTCAAAAAA AGATTCGTAT
 981 ATATCATGTC ATTTGGCATT CAATTCTTT TCACAATTCA TAGTCTATT CTCAAAAAAT TCGAGTCTC
 1051 GTATTGTTAA GGAAAGATCG AAGAGACTGT TCACACAGGT ACTCTCTTTT ATTTATTGAT TCACATTCA
 1121 ATATGTTATT GTTTCTCTGC TTAATGGTT CGTCAGCTCA ACTGCGCTG CTGATTTAAA TTCTCTCACT
 1191 TTCTTCACG GATTTTTAAAT ATTATGGTT TTGTAATGAA CAATTGGTGA AGGAAGAAAA CATGGGAGTC
 1261 TTTCCTAAAG TAAACCTAGA TACTTAGGTT ATAAGGGTAT ATGCTAAAT GAACTATGCC CATTCACTT
 1331 TGCCCTTCTC TTACTTTTGT AGTTTTAGA ATCCAAGTT TCATATGTAT CTGATGTGT GAGAAGAATA
 1401 GGCATTAAGA AGGTAAGGCG CGTACATAAA ATIGATAAT TAGTGAATG TCTTGTAT CATTATTTT
 1471 ACTCTCATAA AAAGCATATA GATCAACAC AAATTGCTAC TTGTTAGTG AACAACCTCG ACTTAAATAAT
 1541 GTTAATAATC AAGATTCCTC TGATTTCAAC TATTCTCAA CGGAACAAGC TCACATTTAA CTCATATTGC
 1611 TTGAGTCG AGTGGTTTAT ATTTGGGGTT TTACATTAA TTTTTGTGC ATGAATGTGA AAATAGACTG
 1681 CTTATTGATT CTGTTGTTG CATTGAGTT ATTTCTTCA TTACTACCTT ACAAAATGCT CAGTGATAGA
 1751 TTTCCTAAAT TTGCTTATT CGGTGCTCT TAAATATGTA GGAGCTACTA AAAGCAAAAAA TATCGAGCAA
 1821 TGTCGGACCC ACAGGGATT GCTGGTGCCTA TTATTAACCC AATTGCTCAG ACGGCCCTGG TTCCCGTTAC
 1891 GGACCATGTA GGCTACATGA TTTCCTGCAG AAAATATGTG AGGGTCATGC AGATGAAAAT GACAGAGTTG
 1961 AATACCTCAA GAATCAGTGT AGAGGAACAC ATTAGCCGA ACACAAAGAAA TCATCTTCAG TTCCATCTCA
 2031 AACTAAGGAA TGGTGGACCG AAGTGAAGG GATCAGAGCA AATGTGGAA ACTTTCCGAT TGATGTCTC
 2101 ACTTGTGTA GTCTCAGGAT CAGGCACAACT TTGGCAGAGA AAGCCTTCAG GATAACTCGAG CAGATTGAAA
 2171 GTCTAACGAG ACAACTCTCC CTGATCAGTT GGACTGATGA TCCAGTTCYT CTAGGAAGAG TTGGTCCAT
 2241 GAATGCACTC ACCCTCTGAC CATTAAAGTGA TGATTTCCCA TCAAGAGAGA AAACCTTTAC ACAAGCACTA
 2311 ATAGCCTCTG AACCCAAACCA AAAATTCCAC ATGGTACCCCT TGTCGGGGAT GGGTGGAGTG GGAAGACTA
 2381 GAATGATGCA AAGGCTGAAG AAGGCTGTTG AAGAAAAGA ATTGTGTTAAT TATATTGTTG GGGCAGTTAT
 2451 AKGGGAAAAG ACGGACCCCT TTGCCCCATCA AGAACCTATA GCAGACATTAC TCGGTATACA ACTCAATGAA
 2521 AAAACTTACG CAGCAAGAGC TGATAAGCTT CGTGAATGGT TCAAAAGAA TTCAAGATGGA GGTAAAGACTA
 2591 AGTTCTCTCAT AGTACTTGC GATGTTTGGC AATTAGTGTG TCTTGAAGAT ATGGGTTAA GTCCCTTTCC
 2661 AAATCAGGTC GTCGACTTCA ATGGCTTGTG GACATCACAG GACTCACAAG TTGCACTPAT GATGGGGGTT
 2731 GAAGCTTATT CAATTATTA CGTGGGCCCTT CTAACTGAG CAGAAGCTCA AAGTCTGTC CAACAATTG
 2801 TAGAAACTTC TGAGCCCGAG CTCCAGAAGA TAGGAGAGGA TATCGTAAGG AAGTGTGGCG TGTCTACCTAT
 2871 TGCCATTTAA ACCATGGCAT GTWCTCTTAG AAATAAAAGA AAGGATGCAT GGAAGGATGC ACTTCTGGC
 2941 ATAGAGCACT ATGACATICA CAATGTTGCG CCCAAAGTCT TTGAAACGAG CTACACCAAT CTCCAAGAAG
 3011 AGGAGACTAA ATCCACTTTT TTATGTTGTC GTTGTGTTCC CGAAGACTTC GATATTCTCA CTGAGGAGTT
 3081 GATGAGGTAT GGATGGGGCT TGAAGCTATT TGATAGAGT TATCGATTAA GAGAAGCAAG AACCAGGCTC
 3151 AACACCTGCA TTGAGCGACT GGTGCAGACA AATTGTTAA TTGAAAGTGA TGATGTGGG TGTCAGAAGA
 3221 TGCATGATCT GTGCGTCTT TTGTTTTGG GTATGTTTC TGAGTCGAG CATGCTCTA TTGTCACCA
 3291 TTGTAATATG CCGGGTGCCTT CTGATGAAAA TGATGATGCT GTGCACTCTT GCAAAAGAAT TTCAATTAA
 3361 TGCAAGGGTA TGATTGAGAT TCCAGTAGAC CTCAAGTTTC CTAAACTAAC GATTTTGAAA CTATGCTATG
 3431 GAGATTAAGTC GCTAAGGTTT CCTCAAGACT TTATGAGG AATGAAAAG CTCCATGTTA TATCATACGA
 3501 TAAAAATGAG TACCCATTC TTCCCTTGGC ACCTCGATGC TCCACCAACA TTGGGTGCT TCATCTCACT
 3571 GAATGTTCAT TAAAGATGT TGATGCTCT TCTATCGAA ATCTATCGAA TCTGGAAGTG CTGAGCTTIC
 3641 CAAATTCTCA CATTGAAATGG TTACCTTCCA CAGTCAGAAA TTGAAAGAAG CTAAAGGTTAC TTGATCTGAG
 3711 ATTTCTGAT GTCTCTCTTA TAGAACAGGG TGTCCTGAAA AGTTTGTCA AACTTGAAGA ATTTTATATT
 3781 GGAGATCTCAT CTGGGTTTAT AGATGATAAC TGCAATGAGA TGGCAGAGCG TTCTTACAAC CTTCCTGCT
 3851 TAGAACTTCG GTCCTTTAAAT AACAAAGGCTG AAGTGAAGAA TATGTCATT GAGAATCTTG AACGATTCAA
 3921 GATCTGAGTG GGATGCTCTT TTGATGAAAA TATCAATATG AGTAGCCACT CATACGAAAAA CATGTTGCAA
 3991 TTGGTGACCA ACAAAAGTGA TGTTAGAC TCTAACTTA ATGGGTTATT TTGAAAACA GAGGTGCTT
 4061 TTTAAAGTGT GCATGGCATG AATGATCTTG AAGATGTGA GGTGAAGTGC ACACATCTTA CTCACTCTC

RLG 2B cont.
SEQ ID NO: 23

4131 TICATTCTGC AATTTAAAAG TTCTTATAT TTCAAAGTGT GTAGAGTTGA GATACTTTT CAAACTCAAT
 4201 CTTGCAAACA CTTTGTCAG ACCTGGACAT CTAGAAGTTT GTGAATGTGA GAATATGGAA GAACTCATAC
 4271 ATACTGGAAAT TGGGGGTGTTG GGAGAAGAGA CAATTACTT CCCTAACGTC AGTTTTTAT CTTTGAGTC
 4341 ACTACCGAAG TTATCAAGTT TGTGCCATAA TGTCACATA ATTGGCTAC CACATCTCGT AGACTTGATA
 4411 CTTAAGGGCA TTCCAGGTTT CACAGTCATT TATCCGCAGA ACAAGTGGC AACATCTAGT TTGTTGAAGG
 4481 AAGGGGTAGA TATATGTTCT TTATGTTAAT ACAATTAAA TAATATTTTC ACCAAATTTC TCATAATATA
 4551 TCTGTAATTG GATTGATGAGA TGTTGTTATG TTATGTTG GCTATTAAGG GATGATTATT TTGCAAGGTTG
 4621 TGATTCTAA GTTGGAGACA CTTCAAGTTG ATGACATGGA GAATATGGAA GAAATATGGC TTGTTGAAC
 4691 TAGTGGAGGT GAGAAAGTTA AGTTGAGAGC GATTAAGTG AGTAGCTGTG ATAAGCTTGT GAATCTATT
 4761 CGCGCAATC CCATGCTCT GTTGATCAT CTIGAAGAGC TTACAGTCGA GAATTGGGGT TCCATTGAGT
 4831 CGTTATCAA CATTGACTTG GATTGTTGCG GTGCAATTGG AGAAGAAGAC ACAAGAGCC TCTTAAGAAG
 4901 CATCACCGTG GAGAATTAG GGAAGCTTAAG AGACCTGTGG AGGATAAAGG GTTCAGATAA CTCCTGATCTC
 4971 ATCAACGGTT TTCAGCTGTG TTCAAGCTAA AGAATTGAAA ATAGTAAGG GTTTGAAGAT ATATTACAC
 5041 CTATCACCGC CAATTATTTAT CIGGAGGCA CTTTGGAGAT TCAGATAGAA GGTGCGGGAG GAAATCAGA
 5111 ATCAGAAGAG CAGGTAACGC TTTCATTTTAC ACTTTCTTAA TTAATTAAGG ACTAAGCTCC TGTTTTTG
 5181 ATAATAAAGA GGTGGGATGA CTAAACTTGG GCATCACAAAT TCCAACAAAA TGTTACAAAC CATGAAACGT
 5251 TCTAACCATTT CTTGAAATTG AGGTTTCAAT ACAAGTCATT TAAAATATG GCTTTAAATTTC TTGTTTATATT
 5321 TATGTTATCAA CATGATTTTT CATTAGAGAT CATTATTATA ATAGTAAGTT TAAAGCAATT TAAATCAGAA
 5391 CTAATCTAA CTITAGCTAA TAAATCGTTA TAAATGTTAA TAAATTCTT TTAGTGAAT AAGCAACGGA
 5461 TTAAATAAGT TAACAACCTA AATGTCATT CCTAACAAAA AAAACTTTGG TTICAGAAAAA CCGCAATTCA
 5531 AGATAACTAA AATAAAATAA TTGACATTC ACTAAGAGCA TTTTTTTTC TAAATATGAT TCCAAATGAA
 5601 TAAAACCTAA ATTATACAG AAAATTCTTT TATATGTT ATACAAATTG TACAATTGAA ATTGGATAT
 5671 GTTAATTAAAC CGTTTATAAT TCTGGTATCA CAAACGGATA TATAATTAA TATTATTTC TGTAGTCATT
 5741 TGTAATTGTA CTAGTTTATA ACCCGTGGGA ACCATGAGTT CTAAAATTAG TTAAACTTTCA ATAATAAAA
 5811 TTAAATAATT TTATTTATTAA TAAATAAAATT ATTAAATTAAG AGATATATCA AAAATTAAA GTTATTATAA
 5881 CTTCAATTAAACATATAAT TAGAAATAT ATGATCATAA CTCTGCACT CTCTGTTAT AAATGCCAGAG
 5951 AAGCTATTAG TATTTCTCA CAAAGTCCA AACCAATTGA AGCTTATATA ATTITGTTGAA AACTCAATT
 6021 GCATTAGTTT TTAAGAGTCA CCAAATTCAA AGAATATTC AATGCTTTCA TTACCACTAT GGAGAAAATA
 6091 TTTCCTTCTAGT TTAAATGAAA TGAACACAAA CATTCAAACCT AATTTGTTGCT TATTAACCCA AAGACCCATT
 6161 ACTTAGCCAA GAGTTAAC AAAAAAAATT ACATCATGT ATCATTATTTC ATGACTAGAT ATATATGAAAC
 6231 ATGAAGGGAG TTTTATAGA AAATATAATC ATAGATTC AACATAACCT CAGGGAAATTCTC CTCAAAATAA
 6301 CCAAGTTATT CAAGAAATTAA CATCCAAGTC AACCAAGAG AAGTTTGGCT TAGCATGGCT AAACCTCAAGA
 6371 AACTAAATA AGGATTAGAA GTACAAACCA TGTAGTAAGA ATCACAGTAA AAGATGATGT TGTTCCTGAT
 6441 GTTCTTCTAA GTTCTTCAG TCTCCAGTTG CTCCCTAAATC TCCAAAGGGAG AGCCATTAAA TTGCTATGTA
 6511 TTGATCCCCCTT CAAAGCTGC ACCAACCTCC CTAAATAAC ACTCAAACCA AAAATGACAA AATGCCCTG
 6581 AGGACCCATT GTGGGTGCTT TGCGGGGTG GAGCTGCATA CGAAAGGTCT TTGGTCTTT TGAGGGTGT
 6651 GTTGTGCGGG ATAGCTTGTG GCATGCTTCC GCGCGGTCA CCACATGTC CACAGTGAT GCATGGTGTG
 6721 TCGCTTCTTG AGTTTTGAGC CTCGGATGCT TAGTCCACTT GGCCCAATTTC GAGTCCAATC AGCTTATAAC
 6791 CCACTTTCTCT TCAACTTATC TTCAAGTTAA GCCCAATTG GCTTCTCCAA ATCATCCATA ACTTCACAGA
 6861 ATCGCCCCGTT CATCTTAAATC CCGGATGCA AATTATTCCTT CGGTCTTCTAT TTAAGCAAG ATACCCACCTT
 6931 CTTCATGCTT CATCCATCAA TAGTACATT CATGATCAT CTCTACTAGT TATTAGTCC ACAAATCCCTT
 7001 GTTGTCTTCC AAATTAAATT ATCTCATTTA GTTCCCCGTT CCGCTACTTCTT CTTAAATT TCGAATTAAAG
 7071 CTCAGAGAAA TATTAAGTAC CGCAAATGGT CATAAAATTAA ACAAAAAGGA AAATGCATGA AGATTAAC
 7141 AATGATGAAC GAAATATGCT AAAATGACT ATAAATGAA TGAATTTATC GCAACTCCGAC
 7211 CACCCCTTATG GCTTGTAGTC CACCCACCCCT CTATCTTGC TACCAATATG GGATGGAAAC ATCATTAAATT
 7281 AAGCCAAAAA GCTAACATAT AAGGGTTAG TGACAAAGGT AAGTACTAA GATGAAAATA ATCCATTTTT
 7351 CTTGTTTTTA CACAAACACAC ACATAGGGC AGACCTAGGA TTTCAGTGTG CAGATTGTTG GTGGCACATA
 7421 ATGTTGCTG GTGACATT TTCTTCTTT TTACGCTGTG GCACAAACAGT AGGAAAAACG AAAAATTGCA
 7491 AATTCTTCTAC AATTGCTTCAAAACCA GGGGGTTG TGCCCAACTG GACAACACAA GTTGAACGTC
 7561 CCTACGCGCG CACACACACA CACACACATA GAGAGAGAGA GAGAGAGAGA AAGAAAGAAA
 7631 GAGAGAGAGA GTTTGGGATG TGATACCTT TTAGGAAAA TGGAGTTATA TCTTGTATAT TGTATTTTT
 7701 TAATGTAATT TATTTTATTAATC ATCATTTAG TTATTAAGTT TATTTTATATIN GNNTATGAAA AAAAAAGTCT
 7771 TTATACATT GGATTTAACAA TAAAATTCGA ACAATTAA TCAAAAGAG CAAACATGTG GACAATTATG
 7841 TATATAATTAA ATTCAACATA GTCTTGTAGA ATAGTCTTAT ATATATAATT AATTCCTCAAT GGTCTTAGGA
 7911 ATAGTAAGTT TTATATTTTC AAACCTTTCG CACAATTCTT TGCTTACTTT GACACTTTTC CTTCTTAAC
 7981 TTACATATAT ATATATTTA AAGGGCAAAG GTCAAGGAA TATAATATTT TCTTATTATC TACGTTTTG
 8051 CACAAAAGTT TGAACACTT GCCACTTTT GTCCCTCCCTT AACCTTTTC AATGTTTGCG ACACAAAGTTC
 8121 CAAACCTTGTG CCACTTTGAT CATTCTCAA CTTTCAACCG CATTAGTTTG TGGAGTTGGC AGTTTGTG
 8191 CCTCTAACTT CGATATTCTC TACTGCTAGC CAAAGGGT TCCAGAGTTT CACACTTTTG GTCCCTGACA

RLG 2B cont.

8261 GTAACCAAT GTGAGATGTC AAATTTTGC CACATTAGTT TGTGGAGTTG TCCCTTTGG TCCCCCACA
 8331 TTGATATTCT TACTATACGA TCTTATTTT CTCAAATAAC AACACGTATA TTTCATC:CT ATTGGAAA
 8401 AGAGTTTTAA AA:AAATAAC GACTAGG::: G:GC:GAGTT TTTTT:ACA AGTTTGATC AAATCATATC
 8471 AAAATTAAG GTGGAACCGGT GACCACATTA ACCAGAAATG TAATTTATTC TTGATTTTG ATAATTTTA
 8541 ATATTTGTT GTGATCTATG TATTTAAAAG TAAACACAA AGAACATAAT CCAAAACCCCT AATTGCAAG
 8611 TCTCGCCCAA TTTCCTATC ACTAGTCCTC ACTTACGATG CGGTAGTC GCCTCTCAC TGCTTACAC
 8681 CCTTGTGTC TACTCATTAC AATAACGAA AGTTGAATAT CCATATATT ATTGGATGT GGAATGAAAC
 8751 GAATCTGGTC AAAATTTGAA TTGTTGTTGAT GGATTGAGT AGAAGTTGG GCAGAACGGG AATGATGGTC
 8821 TGCAGGTGT TATAAACCTG ATTACTATAT ATGTAACCTC TTACAAACGA CCAAGGTTTC
 8891 TTCCAGGTAC CATTGATCT TTTTAACTG TAGTTTCTG AAACACCTG ATTGGATCA ATATCACCA
 8961 ACAACTCTTA AAAACTTGTAT TAATCAATTTT TTTCCTTCAT CTGATAACAA AGTGGAAATGA TTTTCTACTT
 9031 AGATTAACCTT GAAAAAAAG GTCCATGTGC GTCTGGTGG A TCTGGTAAT GAAGATGGAA GGGAGAGCTG
 9101 ACITTAAGA CACAAACACG TCACCATATC TCTTATTTTA TTGTTAAATTG GCTTTGGTGT TATTTCTTT
 9171 TTTCCTTCTT CTTCCTTCT GTATCTCCAG ATGGTATGTC GTGTTGATAA TTACACCTA GAGATTGGGA
 9241 ACAGATGGAA GGGGCTGTG ATTATGGCT CGCCGAGTTT TACTTATTA CTCAATTTC ACCTAAATTTC
 9311 TGATTCTTGT TTGAAAATAA GTGCACTT TATTTTGTG TTACTCTGTT GCATAGGATC CTAGCATCT
 9381 TTGATAAATT TATTGAGG TGAAAGATCC AACTATTTT TAGCTGTTGG CATTTCCAT CATTGCAAC
 9451 TGTTCTTGA AAAAAAAATC CCTAAAATAA AAATAACATC TTCAAAATCC AAAATTATAA GAGAGAATTG
 9521 TAAATGGACA TGGAACTATA AATCATTAAC ACAGTTCAGT AAACAAGTTG CTAATTACAT TTCTGCTGT
 9591 CCAGATTGAA ATTCTATCAG AGAAAGAGAC ATTCAAGAA GCCACTGGCA GTATTTCAAA TCTTGATTC
 9661 CCATCTGTC TCATGCACTC TTTTCATAAC CTCCGTGTC TTACATGGAA TAAATGAA GGAGTGGAGG
 9731 TGTTATTGAA GATAGAGAGT GAGAGTCCTA CATGTAGAGA ATTGTAACA ACTCGAATA ACCAACACA
 9801 GCCTATTATA CTTCCTTACCC TCCAGGATT GTATCTAAGG AATATGGACA ACACGGAGTC TGTTGGAAG
 9871 TGCAGCAACT GGAATAAATT CTTCCTACTT CAAACAAAC AATCAGAATC CCCATTCCAC AACCTCACAA
 9941 CCATAAAATAT TCTTAATGCA AAAGCATTAA AGTACTGTT TTCCGCTCTC ATGCCAGAAC TTCTTCCAA
 10011 CCTAAAGGAT ATCCGGATAA GTGAGTGTGAA TGGTATTTAA GAAGTTGTTT CAAACAGAGA TGATGAGGAT
 10081 GAAGAAATGA CTACATTTAC ATCTTACCCAC ACAACACCA CTTTGTGTC TAGTCTTGTAT TCTCTCACTC
 10151 TAAGTTCTG GGAGAAATCTG AAGTGTATTG GTGGAAGTGG TCCCAAGGAT GAGGGGAGCA ATGAAATATC
 10221 TTTCATAAAT ACCACTGCAA CTACTGCTGT TCTTGTCAA TTGAAAGTAT GCTTTGTACA TATTCCATTAA
 10291 TTATTTTAAT TTCTTTTTT ATTTGCAATA TTCTATAAAT AATACATTTT ATACCCACTA TACTAAGATA
 10361 ATAAATTACCT AGAGGGATGG ATGCTATGAC ACAGCTGCTA CACTCAGAA ACTCTARTAA GGGCAGTTAT
 10431 GGAAGTTCAA TAAAATGATA ATGGCATTCTT TTGATGGTA ATAGGCAA TTAAAGTTT ATTCTGTTA
 10501 AACCGAGTATT TAGCAAGTAC TGCCCGAGTAG GAGAGGAGAA TATCACCTT TGTTGAAATC TGGTCATTGT
 10571 ACCCAGAATT TAGTTAAATG TAACATTITA GATAATTAGGG TTATCAGGT GACAGATATT GTAGAATAGA
 10641 ACACATATGTA ATATTAACCA AAACATTTT TTCTAAGTT GCTCTGTTAA ATATGTGCTT TTCTGATTTC
 10711 ATGAAATTG CATTCTTATA TTGTTAGGTTG TAAAGTATTG TCTCTTCAA TAAATCCCGA AATTTTTTAA
 10781 TTAAAAAAA AAAAAACAAAGTAAATTG TGATATGGAG AGCAGTGGTA TCAATTAGTA TATAAAAAC
 10851 AGATTTGAA TTAAAGTTCT TATATAAAAG CTGTGTATAT AGTTTAAATTG TTTCATCATT TTTTCCAT
 10921 GTGGTGTGTC AGTTGCTGAA AGCAGGTGGT GTTTCTTGGA GCTTATGCCA ATACGCTAGA GAGATAAAA
 10991 TAGGCAACTG CCATGCACTG TCAAGTGTGA TTCCATGTTA TGCAAGCAGTA CAAATGCCAGA AAGCTT

SEQ ID No: 23

RLG 2 B a.m.

MSDPTGIAGAIINPIAQTAALPVTDHGYMISCRKYVRVMQMKMTELNTSRISVEEHISRNRNHLQIP
SQTKEWLDQVEGIRANVENFPIDVITCCSLRIRHKLQKAFKITEQIESLTSQLSLISWTDDPV?LGRVG
SMNASTSASLSDDFPSREKTFQALIALEPNQKFHMVALCGMGGVGKTRMMQRLKKA?EKKLFNYIV
GAVI?EKTDFAIQEAIADYLGQLNEKTPARADKLREWFKKNSDGGKTKFLIVLDDVWQLVDLEDIGL
SPFPNQGVDFKVLTSRDSQVCTMMGVEANSIINVGLTEAEAQSLFQQFVETSEPELQKIGEDIVRK
CGLPIAIKTMAC?LRNKRKDACKDALSRIEHYDIHNVAPKVFTSYHNLQEEETKSTFLMCGLFPEDFDI
PTEELMRYGWGLKLFDRVYTIREARTRLNCTIERLVQTNLLIESDDVGCVKMHDLVRAFVLGMFSEVEH
ASIVNHGNMPGWPDENMDIVHSCKRISLTCKGMIEIPVDLKFPKLTLKLMHGDKSLRFPQDFYEGMEKL
HVISYDKMKYPPLLAPRCSTNIRVLHLTECSLKMFDCCSIGNLSNLEVLSFANSHIEWLPSTVRNLKKL
RLLDLRFDGLRIEQGVLKSFVKLEEFYIGDASGFDDNCNEMAERSYNLSALEFAFFNNKAEVKNMSFE
NLERFKISVGCSFDENINMSSHSYENMLQLVTNKGDVLDKSLNGLFLKTEVLFLSVHGMNDLEDVEVKS
THPTQSSFCNLKVIISKCVELRYLFKLNLANLTSRLEHLEVCECENMEELIHTGIGGCGEETITFPKLKF
LSLSQLPKLSSLCHNVNIIGLPHLVDLILKGIPGFTVIYPQNKLRTSSLKEGVVIPKLETLQIDDMENLEE
IWPCELSGGEKVKLRAIKVSSCDKLVNLFPRNPMSLLHHLEELTVENCGSIESLFNIDDCVGAIGEEDN
KSLLRSINVENLGKLREVWRIKGADNSDLINGFQAVESIKIEKCKRFRNIFTPTANFYLEALLEIQC
GNHESEEQVTLSISLS

SEQ ID NO: 24

SEQ ID NO:

<pre> GGAGAC-ACATATGTCG-AGCCTGAAGAGCTTGCGAGAAGAAAGAAATGTTTATATTTTGTGTTAGCCCGTTATAGCCGAAAGCACGCC → 25 10 20 30 40 50 60 70 80 90 100 RL21A -GGCGAAGACACATAATGATTCCTCCTAAAGACGTTTAAAGACGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 99 - 27 RL21B TGGGGAGACTAGTAAATGTCGAAAGCTGTTGAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 100 - 28 RL21C ---AACAC-ACGTTGCGAAAGCTGTTGAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 92 - 29 RL21D ---AGAG-TCGCCAAACATAGGGAGGTGTTGAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 70 - 30 RL21E ---GGAGAC-ACATATGTCG-AGCCTGAAGAGCTTGCGAGAAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 95 - 31 RL21F CTTCGCGAAAGACACATAATGATTCCTCCTAAAGACGTTTAAAGACGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 100 - 32 RL21G ---GGAGAC-ACGTTGCGAAAGCTGTTGAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 95 - 33 RL21H ---TG-AGAGCTGGAGAAAGCTGTTGAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 73 - 34 RL21I ---AT-GCGA- ---GAAGACCTGGAGAAAGCTGTTGAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 87 - 35 RL21J ---GAAGA- ---GAAGACCTGGAGAAAGCTGTTGAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 71 - 36 RL21K TTGGAAAGAC-ACGTTGCGAAAGCTGTTGAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 99 - 37 RL21L ---G- ---GCTGAGAAG-CCTGAGAAAGCTGTTGAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 45 - 38 RL21M ---GCTGAGAAG-CCTGAGAAAGCTGTTGAGAAAGAAATGTTTATATTTTGTGTTAGCCGAAAGCACGCC 74 - 39 </pre>	<pre> CATTCGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 110 120 130 140 150 160 170 180 190 200 RL22A CATTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 199 RL22B CTTCGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 200 RL22C CATTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 192 RL22D CATTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 170 RL22E ACTTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 195 RL22F CGTTCGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 200 RL22G TATTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 195 RL22H TATTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 173 RL22I TATTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 187 RL22J TATTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 171 RL22K CATTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 199 RL22L CATTGCTATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 145 RL22M CTTTGCGATTGCGCAAGCTGTTAGCGAGTTACCTCTGTTAGCGCTTAAGAAACCTAAAGCAAGCTTAAGCGAGACTCTATAGCTTGCTAAAGCTCAAGGCC 174 </pre>
---	---

ACGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	
410	420	430	440
ACAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	487
RLG2A	RLG2B	RLG2C	RLG2D
ACAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	485
ACAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	477
RLG2E	RLG2F	RLG2G	RLG2H
ACAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	455
ACAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	480
ACGTGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	485
TCTAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	489
RLG2I	RLG2J	RLG2K	RLG2L
TCTAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	467
TCTAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	481
RLG2M			
ACAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	464
ACAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	484
ACAGAGAACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	430
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	
510	520	530	540
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	586
RLG2A	RLG2B	RLG2C	RLG2D
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	584
RLG2E	RLG2F	RLG2G	RLG2H
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	576
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	554
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	579
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	584
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	588
RLG2I	RLG2J	RLG2K	RLG2L
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	566
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	580
RLG2M			
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	564
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	583
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	529
TCTTG-CGGCTTCCATTCACCAAACTTTCACCAATTGTTGAAACTC-----		-----TGACCCGGACCTCA-TAAGATGGAGAGATTTGTAAGGAG	558

GACATTTGAGTC	---	TTCGCCCTGAGT	TTAAACCCGACTCTCA	CAATCTCCAGAGCGAGGAGCTAACTCATTTCCTTGCTGCTTGT	
RLG2A	610	610	610	650	610
RLG2B				660	670
RLG2C				680	690
RLG2D				700	
RLG2E					
RLG2F					
RLG2G					
RLG2H					
RLG2I					
RLG2J					
RLG2K					
RLG2L					
RLG2M					
TTCCTGAGACTTGTATCCCTACTGAGCAGTCATGAGCTGTTGAA	TTT	TGAGTTGAGT	TTT	GGAACTTGTGAA	
RLG2A	710	710	710	750	760
RLG2B				770	780
RLG2C				790	800
RLG2D					
RLG2E					
RLG2F					
RLG2G					
RLG2H					
RLG2I					
RLG2J					
RLG2K					
RLG2L					
RLG2M					
TTCCTGAGACTTGTATCCCTACTGAGCAGTCATGAGCTGTTGAA	TTT	TGAGTTGAGT	TTT	GGAACTTGTGAA	
RLG2A	710	710	710	750	760
RLG2B				770	780
RLG2C				790	800
RLG2D					
RLG2E					
RLG2F					
RLG2G					
RLG2H					
RLG2I					
RLG2J					
RLG2K					
RLG2L					
RLG2M					

GKSEVYEAISVNHN--MPCPPEEND-TVHSCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS					
RLG2A protein DMYSKVEIAISVNHN--TLEPHAN--MDSCKBLSLCKGSEKFPTDIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	310	320	330	340	350
RLG2B protein GMSFEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	360	370	380	390	400
RLG2C protein GMSEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	391				
RLG2D protein GHSEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	399				
RLG2E protein VHSEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	381				
RLG2F protein GMSEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	387				
RLG2G protein GHSEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	389				
RLG2H protein HIFSEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	394				
RLG2I protein HIFSEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	386				
RLG2J protein VSEKLEENH-SIYSCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	391				
RLG2K protein DTRPRPKHSLVYQASVNHN--MSEPKEND-TNSCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	386				
RLG2L protein GMSEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	392				
RLG2M protein GMSEVYEAISVNHN--MPCPDENDKLVISCKRSLISTCKGSEEFPMIAFPNLITLKGMDKSRLPDTYEGKLOVSYDKRKPPLPSPOCS	371				

TNEVRLAH-HDSSLRNFDCSITORTNLLEVLSFANSIEMPLSPTGNIKQLRLLDNYCGLRTENGVNLVKLFLYIGNA-S-FG-----					
410	420	430	440	450	460
RLG2A protein	VANLVLVHIIURCLVWHLKSCQUN-LN-LAVISFANLAVLLEPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2B protein	TNIRVLIILTECSLKHDFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2C protein	TNIRVLIILTECSLKHDFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2D protein	TNIRVLIILTECSLKHDFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2E protein	TNIRVLIILHRCSCMMLMFDCSCIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2F protein	TNIRVLIILHRCSCMMLMFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2G protein	TNIRVLIILHRCSCMMLMFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2H protein	TNIRVLIILHJCSCMMLMFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2I protein	TNIRVLIILHJCSCMMLMFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2J protein	TNIRVLIILHJCSCMMLMFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2K protein	TNIRVLIILHJCSCMMLMFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2L protein	TNIRVLIILHJCSCMMLMFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490
RLG2M protein	TNIRVLIILHJCSCMMLMFDCSSIGNSIEMPLSPTRNLAKLWVYGLVLLVNUKLVLAHLYMIVV	470	480	490	490

SEQ FONO:

SEQ ID NO:

	810	820
AC15-2A	TAGTACTGTTTACCTCTAGA - 56	779
AC15-2B	TAGTACTGTTTACCTCTAGA - 57	777
AC15-2C	TAGTACTGTTTACCTCTAGA - 58	777
AC15-2D	TAGTACTGTTTACCTCTAGA - 59	798
AC15-2E	TAGTACTGTTTACCTCTAGA - 60	721
AC15-2F	TAGTACTGTTTACCTCTAGA - 61	781
AC15-2H	TAGTACTGTTTACCTCTAGA - 62	738
AC15-2I	TAGTACTGTTTACCTCTAGA - 63	722
AC15-2J	TAGTACTGTTTACCTCTAC - C4	784
AC15-2L	TAGTACTGTTTACCTCTAC - C5	699
AC15-2N	TAGTACTGTTTACCTCTAGA - 66	778
AC15-2O	TAGTACTGTTTACCTCTAGA - 67	763

{ } { }

SEQ ID NO:68

RLG3 (real RLG3)
(Strand)

1 AATGCCAAA GAAGTCGGAG CAAGAGCTAA GTTAGAGCAT CTATTTCGAG TCATTATCAT GGTAGATGTC
71 ACTCAAGCAC CCAACAAGAA CACAATTCAA AGTAGTATTG CAGAACAGTT GGGATTTAAA CTGCAAGAAG
141 AGAGCTTGTT GTTAAGAGCA GCTAGGGTAA GTGCGAGGTT AAAATGCTT ACAAGGGTGC TGGTGATATT
211 AGACGGATAA TGGTCAAGGC TTGACATGGA GGAACCTGGG ATTCCCTTIG GATCAGATAG AACACACCAC
281 GGCYGCAAA TCTTGTGAC TTCAAGAAGT ATTAGTGCTT GTAACCAAGAT GAGAGCTGAT AGAATCTTTA
351 AAATACGGAGA AATGCCACTG AATGAAGGAT CGCTTCTTT CGAAAGAAC A GCTAAAAAAG CTCCGAATCT
421 GCATCAAGTA GCAAGAGATA TCGTGGAGGA GTGTGGTGGG C

RLG4
SEQ ID NO:69

1 GAATTCGGTGC TTGGTAAGAC AACTCTTGCC TCTTCGTGTT ATGATGAAAT CTCTAGCAAG TTTGATGGTT
71 GCTGCTTCT AAAAATATCT GGAGGGAATC AAGTAATAAA GACCGTATAG AAAGATTGCA AGAAAAAATC
141 ATTTGTGATG TTTTGAACAC AGACCAAGTG GGCGTAGGGA GAGTTGAAGA AGGAAGGCC ATGATAAAGG
211 ATAGGTACAC ACATAGAAAG GTATTGATTG TGCTTGATGA TGTCGACAC AC TTGAGGCRGC TAGCTAGAAC
281 AGTTGGCTGG ATCACATGAT TGGTTGGTG AAGCTAGCC CATAATAATC ACAACTAGAG ATGACATGT
351 ATTAATTGCA CACAAGTAG ATGTGATACA CAATATAAGG TTGTTAAACA ACCATGAAAC TATGCACTTC
421 TTCTGCAGC AAGCACCCACG GGGTCACAAA CGTATACAG ATTATGACCA ACTTTAAAAA CATGTGGTTT
491 CTTATGCTGG TGGGCTTCCA CTAGCACTGT CGAC

SEQ.ID NO: 70
 RLGI-E169
 [Strand]

1 ATCGTAACCG TTGCTACCGAG ANGCCTGTCC CTCCCTTCATC TTTTGTCTATA TGTCATATTTC TCATNNNATTN
 71 TGGCACATTT AATTTCTGG TTATTTAAA TTAAATTATAA TTCCACATGT CATTTTATGA GTTTTCTAT
 141 TTATTTGAGT TTCACTATAAT ATTAAATGT AATAACANTA AATGCTAT TATTTTCTT TAATAAAACG
 211 CATATAATAT ATAGATTAAA ATCATATAAT ACATAGGTAA AACTCTATA ATACATAATGT TCATCCCCAG
 281 TTATTTATA TGCTTCATCC TTAAATTAT TTTATTTAT TTATTTAGACT AGATGATCTT TGTGATATTA
 351 AAATTTAAT TTGTTCAAAT TTAAATTATA TTAAATTC CACATTGAA ATAAATTTGA ATAAATTTAA AAAAATGGN
 421 CCCACCTTA GTCCATCACT TTTCAGCTC ATCAATATCG TGAGTTTCTC CCTCTGTTTC CACCTTAATC
 491 AATTTTCCA CGGAATGACA GACTCTTAGC CGGTTCCTGAA ATTTCGGTTG CGACACTGTT CATGAAAGGA
 561 GATAATAAT CAAATGGAGC TGCTTCCTG TTCACTTGAG ATGAAAGGTG AATTTGATGT GAAGANAATG
 631 TCAGCGATCN ATCTCCATCC GGAACCCACC ACATPATCG TGTCACCAAC AACACTCAA AACGGYGGAA
 701 GTAGGRAKAC WRKAAGATCA TGAAGAATAG ATTATTTTG TGCTCATGGG CTGACTGAGG AGCGGGTTTA
 771 GTTCATCTTCTTCTTCTG CAAAGAATTA TGCGTCCATC GAATTTTTAC ATCGACAAAG AGTTTCACT
 841 TCGCAATTTT TTGTTTAAACA ATTTTTAACTC TTTCCTTCTT TTGCTGTTAAA CCTCTCAATT GCAACTTGCA
 911 ACTGCAACT TTGGGGCCCA CAAATTTGTCG GTGGGGCTTA ATTAAATCCA CTTATTCCTA GTAAACAATA
 981 ATTCAAAATCC ATCTCTGTC ATCACTTCA TCAACATCTC TTGATAATTTG AAATCACTCA CGCTTCATCC
 1051 ATTTCATCTCA CTCATCTACT ATATTCCTG TTCTTATCAT ATTAAACGAT GCCTGAAATC GTCTTTCTG
 1121 CCTCTCTTGAC AGTGTGTTG GAAAGCTGG CATTTGAAACG CTGAAAGAAG ATTTGTCGCT CAAAAGAAT
 1191 TGAATCTGAG CTTAGAAGAT TGAAAGGAG ATTAGACCAA ATCCAGATC TGCTTAACCGA TGCTTCCCG
 1261 AAGGAACTTAA CTATGAAACG CGTTAAAAAGA TTGCTGTAATG ATCTCCAAAC TTTGGCTTAT GACATAGACG
 1331 ACCTACTTCA TGATTTTGCA ACTGAAAGCTG TTCAWGTTGA GTTGACCGAG GAGGGTGGAG CCTCTTCCAG
 1401 TATGTTAAGA AAACAAATTC CAAGTGTGAG CACAAGTTG TCAACAAAGT ATAGGATGCA TGCAAGTTA
 1471 GATGATATTC CCACCAAGTTT CAAAGAACTG CTGAGGAGCAA AAAAATTTGT TTGTTTAAGT GTGATAAACAT
 1541 ATGAAAAGCC AAAATTTGAA AGGTATGAGG CGTCTTTGGT AGATGAAAGC GGTAATGTCG GACTGIGAAGA
 1611 TGATAGAGA AAATTCCTGG AGAAGCTGGT GGGGGATAAA GATGAACTAG TGCTTAACCGA TGCTTCCCG
 1681 GTGCCCTATG TTGTTATGGG TGGAGTTGGT AAAACAACCTC TAGCTGAGCT TTGTTATGAT GAAAAGAAG
 1751 TGAAGGATCA CTGCGAACACT AGGGCTTGGG TTGTTGTTTC TGATGAGTTG AGTGTCTCCA ATATAGGAG
 1821 AGTTATTTAT CAATCTGTA CTGGGGAAA GAAGGAGTTT GAAGACTTAA ATCTCTTCA AGAGCTCTT
 1891 AAAGAGAACC TTAGGAACCA GCTTCTTCTA ATAGTTTGG ATGATGTTG TGCTGAAAGC TATGGTGAATT
 1961 GGGGAAGAATC AGTGGGCCA TTCTCTGGG CGTCTCTCTG AAGTGAATAA ATCATGACAA CTGGAAAGGA
 2031 CCAATTGCTC AGAAAGCTGG CCTTTCTCA TCAAGACCCCT CTGAGGGTC TATCACAAGA TGATGCTTTC
 2101 TCTTGTCTT CTCAACACCC ATTGCTGTA CCAAACCTTIG ATTCACTATCC AACACTPAAGG CCACTGGAG
 2171 AACITTTTGT GAAGAAATCTG GATGCCCTTAAC CTCCTACGTTT AAGAACACCTT GGAAGGTTAT TAAGGACAAA
 2241 AACAGACAGG GAAACATGGA AGGAGCTGGT GGATAGTGG ATATGGAGGT TAGGAAGAGG CGATGAGATT
 2311 GTTCCGGCTC TTAGACTAAG CTACATGAT CTTCCTGCCW CTGTTAGGCT RTRRTTTGCA TAYTGTCTCT
 2381 TGTTCCTCCA GAAGTATGAC TTGACAAGG AGGAGTGTG TCTATGTTGG ATGGCAGAAG GGTTTTGC
 2451 CCAACCAACTT AAAYAAACATC CAAAGCAACG KTTGGCTCTT GAATTTTTR AAGAGTTTT GTCAAGRTR
 2521 TTTTTCAC ATGCTCCCTAA TRRCAAAATCS TTGTTGTTGA TGCAATGACCT AATGAATGAT TTGCTCAT
 2591 TTGTTGCTGG AGAATTTTGT TCAAGGTTAG ACATAGAGAT GAAGAAGGAA TTAGGATGSS AATCTTTGGA
 2661 RAAGGACCCG CATACTGCTA TTGTTGTA GRATTACATA GCTTACAAA RCTTCGAGCC ATTTCAGAGGA
 2731 GCTAAAAAATT TGAGAACATT TTTCATCTG TTGTTGGGG TGTTAGAAGA TTGGAAGATG TTTCATCTAT
 2801 CAAACAGGT CTGAAATGAC WTACTTCARG ATTACCAATG GTTAAGGGTC CTRAKTTTCA TTRRTCTTAY
 2871 AATAASYR-S TGACCAAAK TGCTGGTAG TATGAAASCAC TTGCGGTATC TTAAATCTTAC WGRAACTTWA
 2941 ATCACMCATT TACCGGAAWA TTCTCTGCRAT TTATTAATTC TACARACCCCT GATTTTCTC GCTCTGAMT
 3011 ATTAGTTTAAT KTGCCCCAA ACCTCTCTAA ASCTTTAAA TTGTCASCAT TTGACATGA GGGRTACTCC
 3081 KAAKTTTRAAR AACATCCCCT TARGGATTCG TGARTGAAA ARTCTACAAA CTCTCTTYMG TAACATTGGC
 3151 ATAGCAATTA CGGACTCTAA GAACCTTGCAC AAYCTCCATG GGAACAAATTG TTGTCGGGG CTGGGAAAAA
 3221 TGGAAAATTCG MGTCGAGTC ACGTTAAGGG AACTTGTCTC A: AAAAAGGT IWAATGARTT ANAAACTGRR
 3291 WTKGGGGGTG ATRAAATTAA TTGTTTCCGA ATGGGAACAA CTGAAAAAAA NAAGGTCTCTC AATGAATTGA
 3361 ATGCCCTCATC ATGGTAYTC AAMWARRRY YWTWAWRWT TWMGKAWRKK GKTTTYATRR TKTMYRAAW
 3431 WAGRGTKTTR KARGTAGGTG TTCACTTCA ACCTTCAACTC ACCCTAGTGG GAAATAGAT GATTTTTCA GGGCTACTG
 3501 ATGAGATCTG GAGGGTATG ATAGGGTIVTC TTGGGGGGT AGAAGAAATA ACCATCCATT TTGTTAATGA
 3571 AATAAGGATAT YTGTGGGAAT CAGAACGACA CGCAAGTAAG TTCTCTTATGA ATTAAAGGA TTGGAATTAA
 3641 GGTGAATCTG AAAATTGGT GAGTTTGGG GAGAAAAAGG AGGATAATCA TAATTTAAT AGTGGGAGCA
 3711 CCTCTACCTC TTTCAGGAG TGATGTTGATG TGAGATGTTG TGAGATGTTG CACCTCTGGAG CTTGCTGAG
 3781 CATGGAGAT TTGTTATGTC ACATCTGTC TTCAATNACA TCCGCTCTCTT TCCCAACAGG AGGAGGACAG
 3851 AAGATCAACT CACTTACCAT CACTGATTGC AAGAAGCTTT CGGAAGAGGA GTTGGGAGGA CGAGGAGAGA
 3921 CAAGAGTGT TATAAATCTA AAAATGCCAGA TGCTTGAATC ACTAGATATA CTAAATTGGC CAAATCTGAA
 3991 ATCTATCACT GAATTTGAGT CCTCTATTC CTCGAACAGA TTATATATAT CAAACTGTCC GAGTRTGGAG
 4061 TCATTTCTG ACCATGAGT CGCAAACTTC ACCTCTTAA CAGATGCAAG GAGGAGACAG CGATTTCTC

RLG1-E169
[Strand]

4131 ACGAACGGTT ACGATTCGAC TGGCCGTGCT TTT

SEQ ID NO: 70

Further Characterization of RG2 Family Members:

Further sequencing of cloned RG2 polynucleotide sequences, as discussed above, identified additional RG2 species, listed below. Additionally, further sequencing of the 5' sections of RG2 sequences listed above resulted in modified and/or new sequence 5 information, also listed below. The AC15 sequences found in the 3' sections of RG2 family have not changed.

Listed below are: four full length species, RG2A, RG2B, RG2C and RG2S; two near complete, but with a gap in the largest intron, RG2D and RG2J; three nearly 10 complete RG2 gene sequences, RG2K, RG2N, and RG2O. The deduced translation products (polypeptides) encoded by these RG2 species are listed below. The polynucleotide sequences do not contain any gaps (as with some of the polynucleotide sequences), because all of the gaps in the sequences are in introns, *i.e.*, there are no gaps in exon, or coding, sequences.

They include: an RG2A polynucleotide sequence (SEQ ID NO:87) and its 15 deduced polypeptide sequence (SEQ ID NO:88); an RG2B polynucleotide sequence (SEQ ID NO:89) and its deduced polypeptide sequence (SEQ ID NO:90); an RG2C polynucleotide sequence (SEQ ID NO:91) and its deduced polypeptide sequence (SEQ ID NO:92); an RG2D polynucleotide sequence (SEQ ID NO:93) and (SEQ ID NO:94), and its 20 deduced polypeptide sequence (SEQ ID NO:95); an RG2E polynucleotide sequence (SEQ ID NO:96) and its deduced polypeptide sequence (SEQ ID NO:97); an RG2F polynucleotide sequence (SEQ ID NO:98) and its deduced polypeptide sequence (SEQ ID NO:99); an RG2G polynucleotide sequence (SEQ ID NO:100) and its deduced polypeptide 25 sequence (SEQ ID NO:101); an RG2H polynucleotide sequence (SEQ ID NO:102) and its deduced polypeptide sequence (SEQ ID NO:103); an RG2I polynucleotide sequence (SEQ ID NO:104) and its deduced polypeptide sequence (SEQ ID NO:105); an RG2J polynucleotide sequence (SEQ ID NO:106) and (SEQ ID NO:107), and its deduced 30 polypeptide sequence (SEQ ID NO:108); an RG2K polynucleotide sequence (SEQ ID NO:109) and (SEQ ID NO:110), and its deduced polypeptide sequence (SEQ ID NO:111); an RG2L polynucleotide sequence (SEQ ID NO:112) and its deduced polypeptide sequence (SEQ ID NO:113); an RG2M polynucleotide sequence (SEQ ID NO:114) and its deduced 35 polypeptide sequence (SEQ ID NO:115); an RG2N polynucleotide sequence (SEQ ID NO:116) and its deduced polypeptide sequence (SEQ ID NO:117); an RG2O

polynucleotide sequence (SEQ ID NO:118) and its deduced polypeptide sequence (SEQ ID NO:119); an RG2P polynucleotide sequence (SEQ ID NO:120) and its deduced polypeptide sequence (SEQ ID NO:121); an RG2Q polynucleotide sequence (SEQ ID NO:122) and its deduced polypeptide sequence (SEQ ID NO:123); RG2S polynucleotide sequence (SEQ ID NO:124) and its deduced polypeptide sequence (SEQ ID NO:125); an RG2T polynucleotide sequence (SEQ ID NO:126) and its deduced polypeptide sequence (SEQ ID NO:127); an RG2U polynucleotide sequence (SEQ ID NO:128) and its deduced polypeptide sequence (SEQ ID NO:129); and RG2V polynucleotide sequence (SEQ ID NO:130) and its deduced polypeptide sequence (SEQ ID NO:131); and, an RG2W polynucleotide sequence (SEQ ID NO:132) and its deduced polypeptide sequence (SEQ ID NO:133).

Characterization of New RG Family Groups and RG Species:

Further BAC insert characterization and sequencing, as discussed above, identified new RG polynucleotide sequences. The new sequences were characterized as belonging to new RG families; designated RG5 and RG7. These RG polynucleotides sequences, and their predicted translation products (the polypeptides which are encoded by these sequences) are summarized and listed below.

Identified and listed below is an RG5 family member, designated as the RG5 polynucleotide sequence set forth in SEQ ID NO:134, and its deduced polypeptide sequence (SEQ ID NO:135). This sequence contains an NBS region sequence.

Also identified and listed below is an RG7 family member, designated as the RG7 polynucleotide sequence set forth in SEQ ID NO:136. No deduced polypeptide sequence is given for the new RG7 family member as this sequence appears to be a pseudogene.

RG2A polynucleotide sequence (SEQ ID NO:87)

AAAGTTCATATCCAAGCTTGCCTCCAACCTCTAGCTCCTTAATGGCACCC
30 TCCTTCTCTTCAAAAGCACACAAGAACACTTCAAGCTCAACCACACTCA
CACAAAGCTCTAGAACGAGGGTTAGGGCACATTAGGGTTTGCTCTCTGG
AAATGGGTGCTAAAAGTGAGGCCATAATGTTCCCTATATAAGGCTCACTC
CCACAATTAGGCTTCAATCTGAACGTANTACGCCAGTGTACACTATGG
TACGCCAACGTACTCGGTAGTCTCCCGTCAANAATACACTCATGAGTA

CGCGCAACGTACTTCCCTTACGCCAGCGTACTCAAAGCAAACATT
TTTCAGGACTAATTTGACAACTTGAGGAAGAAAAGGATCAAAGANA
TATACTTGAATTCCGGATGTTACAATGAAGTGAGTGANACCTGGCTAAAAA
ATTAATTGGTTGTGGAAGCCGTGGCTGAGCAAGCAACAAGGGTAAAAT
5 TCGTAATCTACAAATGGTGTATTTCTATTCCTTATTATTTACTT
GATTACGGTAGTTTTCTTACAAAAAATTTAAAGTTGATAAG
TATGCCACTAAAATTGACTTTCCAAAACATAATGTCAAATGGTGC
ATATGTATCATGTTTATTANATAATGAATATGATGATNCTGTTCTATT
AANCGAAAAAATTATCTAATGATTATATTGGAAAACAAAGTTGTGAT
10 TTTNGCATAATATAATCAAATCCNCTTGTNTGGGAGGTGGATAATG
TGGTAAATTANAACAAGTGTNTACNTGAAAGGTNTGAAAGGTGA
AAAAAGTTAAAATGATAAAATGTTACACAAATGTTGATCCGACTGAAT
ATNATGTTAAGGATNATTGATTAAATTGTTGATATATAGTAAGCATAA
ATATTAGAATTGACTAAATTATAAGTTATNCNACTGGATTGAAA
15 CATTGGATATANATTAGGAATGAGAAAATGAGCAACCCTAACATACTTAT
CTTGGTAGTTGGTTATTATATTATTANAAATAGAACATCCCTT
TATTAAACCCATTGAGCGACTGAAATAATGGAAAAATGTAC
CTTGCTATTAGCACAAAAAAATTATAAAATGTCATTGCTATTAGCA
CAAACAAAAAAACTATCCTTTGCATTAGTCACAAAGAAATA
20 TAAAATGGAAATGTTGCTATTAAATGCACTAAAGAAACTATTTGC
CTTATTAAACCGGGTAAACCAATAGAAAAATGGAAGTACATTGTCATT
AGCAGAAAAAAATAACTTCCATTGGCATCCGGTCACAATAATAG
AAAAATGAAAGTACGTTGCTATTAGCGAAACTAACTCCTTTCTT
TTGGCATCGTATCATAAAATAGACTAAATACGTTAGTTACATT
25 TAATACATTGAAATGTCTAATCCACATGTTATTCTATAAAAGGGAAATG
TAATTACTTATTCTTGATTCTTGGCTTCTTTAGTACCCAAAACAT
CCCTCTATCCATCTATTCAACTAAATAATGAAAACATATTCCCTCCA
TTGTTAGGGATGTTATAAATTGTAATTGTTTATGCAAAAAAGTGT
TTGTTAACTAGATTAACGAGATTCACTTTCAGCATTAGGAGAAGTT
30 CATCCATCTTGATATGAAGTGCAAGCCAAGTCTTAACATGGAATA
TGAGGTCCCTATATGCTAAAAAATGCAAATGAGAAATTGTTAAATTG
GATCCCCATAAAAGAAAATTGTTAATGGTTTTAATATTGGTCAATG
TGTCCACCGGATGAGCATAATACTAGTTATAAGGGTAAAGGTGGTT
GGTGGGCCATTATCTTATTATTCTAAAGTCAGAATTAGTAAAAA
35 AAATTATAAGATAAAACCATAACGATAAAAAATCATTATTGGACCA
AAGACCAAAGTGTAAAGGGCTGTTGTTTTGTGAAGAGCTGTGC
AACCACTTTGTCGCGCCGACAGACAACGTGCAGACATATGCCCTCGC
AGAGTGTGTTGTTTGAAAGTGCAGACACAAAAACGTCGCGAG
GTCATCCTGGCGCATATATGTCAGTCTCAAAGGTCTCAGACCTC
40 ATTTAACCAAAAAAAAGACCACCGTTTTTTTTNTTC
TTTCTTTGTTAGCTGAAATGCATTAAATCTTATGACATGAAATTAA
GTTGAAAAATTAAATTATTCAACAGCTGTAGACGTTAAAACAAACAG
TCTTCTTGTGCAGACTGTGGACATTGGCCACCTCTACCGCAGAG

ACTTGCAGATGTGGTCCGAGACTGCAGACATTGGCTCAAATAACA
AACATCACCTAATTGACTACACCACACGGACCTCCAATGTAACAAAAAA
AAGGTTGAAACAAAGTTGCCTATTCTCCATATCCAGGGGCCATTATGT
AAGAGTTATCTAAATTTAGTCGGTAGATCAGTTCTCACATTTAACCG
5 GGTAAAGTGTATGTGTACCGCGCACCTGAAAGGTTGAANGTAACCT
CCA-AACTGAANCAANAATCGATATGAAGTATCAAGTTAGAGGTTCAATTG
GTGAAGGAATCAGCTGGAGGTTGGGAATCGAGCTTCACTATTAAGGTA
AAATCCATAACCCCTAAATGTTGGTACGCTCATATATCAAATTGCGTGT
TGTGAATGAAAAAAGCATGCTAAAAACCAAGTGTAAAGGCACGGTATAT
10 GACATATTATAGTTACTGATAACAAATTATGATAATTGGTTACGT
AAGTTAGGATTCTGACTTCACCAAATGTAATAGTTTGAGTCTATC
TATGTATTGGGAATCACATTAGCAACGGATTGTACTAGTAATTGAA
AAAGTCTTTAAATAATTCTGTITATAATTATGAATAGTTAGCG
ACATCTAATATTAAATAGAATGTATCTGATATTGAATTAATGCTCTTAAT
15 GTGAACATAGACCTTCCATTACTAATGCCTAATTATTAGTTCTAAT
CAATAAATTAAATTCTGTTTATGCTTCTAAGACAATAAAATCCATG
ATTACCTTAAATATTAACAAAATGACCATAAAATAAATAAAATTAG
GATACCAAACCCCCCGCCATGCCAATGTCTAAATATTCTGATGCTT
TGCTTTCCCTCTTCTGTAGTCTATTATTCTGGAGAGTTGAGAG
20 AGTTTCATACAAGAAAATTCAAGAAGAAAGCAAAGGTCCAGGTATTCTC
TTTCTTAATTATGTATTAACCTACAAGCATTTCACACGATCCATGGT
TTTTGTGTATGTTTCAAATTGAAACTAGATTGGACTTTGCCCTG
ATGATTCTAAGATATTGCATGGAGTTGAGATTGTGAAGAAAAGTGGT
AATAGAAAGAGCAAGTGAATCCAGATATAGTATTGGTAATATGATGAT
25 GAGATAGAGATATGTTAAAACGGCTAGAAAATTGTTAATTGAAATT
TAGGTTGTTGAATTGAAAGATACCAAGCTAATAACTAATTAGTTATGCT
AAAATAGTTATAAGAACACAACAAACTCGTAGTTTTTCATGATTTCA
ACCTCTCGTACCAAACAAATTATAACAAAATTGAATATCATTCTCTGC
AATCAATTAACTTTGTTATTATCATCATGTCTAAATTGCCACAAGT
30 TTATTTCATAGTCATATTGGATTATGAAAGGACTATTTCACCAATTAC
ATCTTACTTTATGCCAAAGCTAATACAATCCGACTAAACTAAAGGATT
CTAGGATGCATATAGTTGCCACCAAGTAAATTCTGAAATGGATGT
CTATTGACTAATTAGGTGCCACCAAGTAAATTCTGAAATGGATGT
CGTTAATGCCATTCTAAACCAAGTGTGAGACTCTCATGGTACCCGTTA
35 AGAAACACATAGGGTACCTCATTCTGCAGGCAATATGAGGGAAATG
GGTATCAAATGAGGGATTGAATGCTACAAGACTTGGTGTCAAGAGCA
CGTGAACCGGAACATAAGCAACCAGCTGAGGTTCCAGCCCAGTCAGGG
GTTGGTTGAAGAAGTAGGAAAGATCAATGCAAAAGTGGAAAATTCCCT
AGCGATGTTGGCAGTTCAATCTAAGGTAGACACGGGTCGGAAA
40 GAGAGCCTCCAAGATAATTGAGGACATCGACAGTGTCTATGAGAGAACACT
CTATCATCTTGGAAATGATCATTCCATTCTTAGGAAGAATTGATTCC
ACGAAAGCATCCACCTCAATACCATCAACCGATCATGATGAGTTCCA
GTC.AAGAGAGCAAACCTCACAGAAGCACTAAACGCACTCGATCTAAC

ACAAATCCCACATGATAGCCTATGGGAATGGCGGAGTGGGAAGACG
ACAATGATGCATCGGCTAAAAAGGTGTGAAAGAAAAGAAAATGTTAA
TTTATAATTGAGCGGTGAGGGAAAAAACAGACCCATTGCTATT
AATCAGCTGTAGCAGATTACCTAGGTATAGAGCTCAATGAAAAACTAAA
5 CCAGCAAGAACTGAGAAGCTCGGAAATGGTTGTGGACAATTCTGGTGG
TAAGAAGATCCTAGTCATACTCGACGATGTATGGCAGTTGTGGATCTGA
ATGATATTGGTTAACGCTTACCAAATCAAGGTGTCGACTCAAGGTG
TTGTTGACATCACGAGACAAAGATGTTGCACTGAGATGGAGCTGAAGT
TAATTCAACTTTAATGTGAAAATGTTAATAGAAACAGAAGCACAAAGTT
10 TATTCCACCAATTATAGAAATTCTGGATGATGTTGATCCTGAGCTCCAT
AATATAGGAGTGAATATTGTAAGGAAGTGTGGGGTCTACCCATTGCCAT
AAAAACCATGGCGTGTACTCTTAGAGGAAAAGCAAGGATGCATGGAAGA
ATGCACITCTCGTTAGAGCACTATGACATTGAAAATATTGTTAATGGA
GTTTTAAAATGAGTTACGACAATCTCAAGATGAGGAGACTAAATCCAC
15 CTTTTGCTTGTGGAATGTATCCGAAGACTTGTATTCTTACCGAGG
AGTTGGTGAGGTATGGATGGGGTTGAAATTATTAAAAAGTGTATACT
ATAGGAGAAGCAAGAACCAAGGCTAACACACATGCATTGAGCGGCTATTCA
TACAAATTGTTGATGGAAGTTGATGTTAGGTGCATCAAGATGCATG
ATCTTGTTCGTGCTTGGATATGTATTCTAAAGTCGAGCATGCT
20 TCCATTGTCACCATACTGAAATACACTAGAGTGGCATGCAGATAATATGCA
CGACTCTTGTAAAAGACTTCACTAACATGCAAGGGTATGTCTAAGTTTC
CTACAGACCTGAAGTTCCAAACCTCTCCATTGAAACTATGCATGAA
GATATATCATTGAGGTTCCAAAAACTTTATGAAGAAATGGAGAAGCT
TGAGGTTATATCCTATGATAAAATGAAATATCCATTGCTTCCCTCATCAC
25 CTCAAATGTTCCGTCAACCTCGCTGTTCATCTACATAATGCTCGTTA
GTGATGTTGACTGCTTTGATTGAAATCTGCGAATCTAGAAGTGCT
TAGCTTGCTGATTCTGCCATTGACCGGTTGCCTCCACAATCGAAAGT
TGAAGAAGCTAAGGCTACTGGATTGACAAATTGTTATGGTGTGCTATA
GATAATGGTGTCTTAAAAAATTGGTCAAACACTGGAGGAGCTATATGAC
30 AGTGGTTGATCGAGGTCGAAAGGCATTAGCCTCACAGATGATAACTGCA
AGGAGATGGCAGAGCGTCAAAAGATATTGATGCTTACAGACTGAGTT
TTGAAAACGATGCTCAACCAAGAATATGTCATTGAGAAGCTACACG
ATTCCAGATCTCAGTGGGCGCTATTATATGGAGATTCCATAAGAGTA
GGCACTCGTATGAAACACATTGAAGTTGGTCTTGAAAAAGGTGAATT
35 TTGGAAGCTCGAATGAACGAGTTGTTAAGAAAACAGAGGTGTTATGTT
AAAGTGTGGGAGATATGAATGATCTGAAAGATATTGAGGTTAAGTCATCCT
CACAACTTCTCAATCTCTCGTCAACAATTAAAGAGTCCTGCGTT
TCAAAGTGTGCAGAGTTGAAACACTCTTCACACCTGGTGTGCAAACAC
40 TTTAAAAAAGCTTGAGCATCTGAAAGTTACAAATGTGATAATATGGAAG
AACTCATACTGAGCAGGGTAGTGAAGAAGAGACGATTACATTCCCCAAG
CTGAAGTTTATCTTGTGGCTACCAAAGCTATCGGGTTGTGCGA
TAATGTCAAAATAATTGAGCTACCACAACTCATGGAGTTGGAACTTGACG
ACATTCCAGGTTCACAGCATATCCCATGAAAAAGTTGAAACATT

AGTTTGTGAAGGAAGAGGTAATATAAATTAAATGCTAATACATTAC
AAAGGATCTTCAGTTAAATCTTCAAAATATTTGATTGATGGT
TGGGGTATTATTGTTGGATGGACTATTAATAATGATTATCTGCAGGT
5 TCTGATTCCAAGTTAGAGAACTGCATGTTAGTAGTATGTGGAATCTGA
AGGAGATATGCCCTGCGAATTAAATATGAGTGAGGAAGTTAACAGA
GAGATTAAAGTGAGTAACGTGATAAGCTGTGAATTGTTCCGCACAA
GCCCATATCTGCTGCATCATCTGAAGAGCTAAAGTCAGAATTGTG
GTTCCATTGAATCGTTATTCAACATCCATTGGATTGTGTTGGTCAACT
GGAGATGAATACAACACAGTGGTGAAGAATTATAAAGTGATCAGTTG
10 TGATAAGCTGTGAATCTCTTCCACACAATCCCAGTCTATACTGCATC
ATCTGAAGAGCTGAGTCGAGAATTGTGGTCCATTGAATCGTTATTG
AACATTGACTTGGATTGTGCTGGTCAATTGGCAAGAAGACAACAGCAT
CAGCTTAAGAAACATCAAAGTGGAGAATTAGGAGCTAAGAGAGGTGT
GGAGGATAAAAGGTGGAGATAACTCTCGTCCCTGTTCATGGTTCAA
15 TCTGTTGAAAGCATAAGGGTACAAAATGTAAGAAGTTAGAAATGTATT
CACACCTACCACCAAATTAAATCTGGGGCACTTTGGAGATTCAA
TAGATGACTGCGGAGAAAACAGGGAAATGACGAATCGGAAGAGAGTAGC
CATGAGCAAGAGCAGGTAGGGATTCAATTCACTGTCTTAATTAAATGAT
TAAGCTCTGTTTGAATAAAAAAGGGACAAACCATTCACTGACTTAA
20 TGTAGCAATACAAGTCATGTATAAGAGTGACCAACTTTTTATTATA
AAATGACTACAAAATATTTCATTAGAGATCATGTATAATGTGAC
TAATTTCATCACCTAACTTAGTTGATAAATCTTATAATGTCACTA
GTTACTTTCAAGAAAATAACAAATTAAATAAAATTCAACAAAAAGCAT
CAACTAAAAAAATCCCACACCCGTAATAATTAAAATAAAAGGATTAA
25 CATCTAATACGAACAATTTCCTAAACATGATTGGACCAATATCA
CCAGCAACTCAAGTTGGAATCGATTCACTGCTAAACTGACCAGCATAA
TTAGATAGATGAGAGTTGAAGCTAAAGTGCCTATATAAGTTGTTCATC
TTTTCTTGATCTGATAGCAAGTTGAATGATTCTTCTTCAAATTG
ATAAAAATCTACATTATAAAGAGACTAGCTGAAAAAAATGGCTAGGT
30 GGGTCTGGTTCTGGTAGATGAAGATGGAAGGGGAGAGTAGATTCAA
GACACAACACATCCTTCATTTTATTATTATTATTATTATTATTG
ATATCTGCTCATTTGTTACAGATATGTGAGGTCTATTAAATTTTA
AATATATAAAAAATAAACATAATGAGAAAATTAAATAAGAATAA
ATTAAATAAGGGCACAATAGCTTTAGGTAAGACAAGGACAAACACGC
35 AACAAAAATAAACAGTAGGGACCATCCGATTAAAAAAATAATTAGGGA
CCAAAAACATAAATTCCCCAAACCATAGGGACCATTCATGTAATTACT
CTTACTTTCTGTTGTTCATTTGGGTAACTATTCTTGTACACAT
CTAGGTAACGAACCTGTTGAAGTGTCCCATTAGGATGTGACCTACTAC
AACCGATCATAATAGTCATGTGAACACTTCAACAACTTTATTACTTA
40 GGTGTACAAAAAAACAATAGTACCATGATGTGAACACATACTGAAAAT
TAATTACCTTAGCAAGTTATTCTTCCATTAGGTTGATGGAAACAGITC
CGTGAGACCGTGACTGGATGGTAGATAAATTAGTAAACTAACCTTC
AATTAAACCTACCTTTCTTATTAACTCAATTCAACCTAAATTCTGATT

CTTGTGAAAGTAAGTTGCATCTTATTGGTATTATCTTGTGCATA
GGATCCTTAGCATCTTAAATAATTATTGAAGGTGAAAGATCCAAC
TTTTAATCTGTGGCATTTCCATCAATTGCAACTGTTCTGAAAAAAA
AAATACCTAAAATCAAATAACCATTCAAAATCCAAATTATAAGAGAG
5 AATTGTAATGGACATGGAATCATAAACATTAACACAGTCAGTAAACA
AGTGCTAATTACATTCTTGTGCAGATTGAAATTCTATCAGAGAAA
GAGACATTACAAGAAGCCACTGACAGTATTCTAATGTTGATTCCCATC
CTGTCTCATGCACTCTTCAACCTCCAGAAACTTATTAAGACAGAG
TTAAGGAGTGGAGGTGGTGGAGATAGAGAGTGAGAGTCCAACAAAGT
10 AGAGAATTGGTAACAACTCACCATAACCAACAAACACCTATTATACTTCC
CAACCTCCAGGAATTGATTCTATGGAATATGGACAACATGAGTCATGTGT
GGAAGTGCAGCACTGGAATAAATTCTTCACTCTCCAAAACAATCA
GAATCCCCATTCCACAAACCTCACAAACCATAAAAATTATGTATTGCAAAG
CATTAAGTACTTGTGCGCTCTGGCAGAACTTCTTCAACCTAA
15 AGCATATCAAGATAAGAGAGTGTGATGGTATTGGAGAAGTGTGTTCAAAC
AGAGATGATGAGGATGAAGAAATGACTACATTACATCTACCCACACAAC
CACCACTTGTCCTAGTCTGATTCTCACTCTAAGTTCCGGAGA
ATCTGAAGTGTATTGGTGGAGGTGGGCCAAGGGATGAGGGGAGCAATGAA
ATACTTTCAATAATACCACTGCAACTACTGCTGTTGATCAATTGAA
20 GGTATGCTTGTACATATTCAATTATTATTAAATTCCCTTTTATTG
CAATATTCTATAAATAATACATTATACCCACTATAACTAAGATAATAAT
TACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTCAGAAACTCT
AGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATTTGAT
GGTAATATAGGCAATTAAAGTTTATTCTGTTAAAGCAGTATTAGCA
25 AGTACTGCCAGTAGGAGAGGAGAATATCACCTTGTGAAAATCTGGTC
ATTGTACCCAGAATTAGTTAAATGTAACATTAGATATCAGGGGTCA
CAGGTGACAGATATTGAGAATAGAACAAATATAATATCACCCAAA
ATTTTTCTAAGGTATTCTGTTAAATATGTGCTTCTGTTCTCATNGA
ATTNGCATTCTGATATTAGGTGTTAAAGTGTATTNTCTCAATAAAAT
30 CCCGAAATTAAATTAAAAAAAAACAAAAGTACATTGATGTGGAG
AGCACTGGTATCACTTAGTATATAAAAGCTGATTTGAATTAACCTTC
TTATACAAAAGTTGTATATAGTTAATTAGTTACATCATTCTCA
TGTGGTGTGCAGTTGTCTGAAGCAGGTGGTCTGGAGCTTATGCC
AATACGCTAGAGAGATGAGAATAGAACATTCTGCAATGCATTGTCAAGTGT
35 ATTCCATGTTATGCAGCAGGACAAATGCAAAGCTCAAGTGTGACAGT
AAAGTATTGCAAAGGGATGAAGGAGGTATTGAAACTCAATTAGGAGGA
GCAGCAACAAAACAACAAAGAGTGGTGCAGGTGAGGAAGGAATTCCAAGA
GTAATAACAATGTTATTATGCTTCTGGTCTGAAGATATTGAAATCAG
CTTTGTGGGGTTGGAACATATATTACATTCTCTGCACCTGAAAGCC
40 TGAGACAGCTCAAAGAGTTAAAGATAACATTGCTACGGAAATGAAAGTG
ATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGTAAACAACAAAC
AACACAAATAACGAAGGGGATCATCATCATCATCTTCAATGAACTAAATGATGTA
AGGAGGTTGTGGTCTTCCTCGTCTCAAATCCATTGAACATGAACTAAATGATGTA

CCAGAGCTGGTAGGATTCTTCTGGGAAGAACATGAGTCGGTTGCCTC
ATTGGAAGAAGTACCATCAAGTATTGCTAAAAATGATGGTGTGCAG
CTGGTGGGTCCACAGCTCCCCACTCAAGTATACACACAGAATTAGGC
AGACATGCTCTTGATCAAGAACATCTGGCCTTAACCTTCATCAGGTATAT
5 ATTTCTTAATTGGCATCATCTAACTAACAGATATCATTGCCAAG
TAAATTACTCAAACACATTACACTGGTTCACTAAGTTATGTTG
TTCTAGGAAGGCCAAATGGGAAGCAAGATAGGGAAAAATAGTGTATT
CAGTGGAAAGGGTATTTAGGTATTTCTGTCAAAAGTTGTTATTGCAGG
10 CTTTTAGTACCTGGAATCGTGTGGGAGGAGCATTATTCTGATT
GCTTGTCTTATCATTCTTAGCCTCTGGAACAGCTAGAAACCCT
TTTAAATCTTGTATTTCAATGACAAAATTTCTGTACTACATTGA
TTGTTGTTCTCATGGTCTAAGTGAGTTATTGGCTCATCTGTTACTTCT
TTTGATTGTTATTTCATATCATGTTAGTCACTGAATCAAGCTTTCTA
TTTCAACCAGGGCAAAAGGTCAAAAGTAACCTACTTATGAGATCAAAA
15 ACAGCAACCCATCGATAACTTTAGTGGAGTTAATAGTTACAATTACC
ATTGTGATTAATAATTATAATCTTGATTAATTCAAAAAATTGGTAC
AGCACATATATGACATTCAAAGGTTTGACATATATATGCCTCT
GGCGTTTCTTATTGGACATGCAGACCTCATTCAAAGTTACGGTG
ACACCTTGGGCCCTGTAACCTCAGAAGGGACAACCTGTTCTTCATAAC
20 TTGATCGAATTATATATGGAATTAAATGATGCTGTTAAAAGATTATTCC
ATCCAGTGAGTTGCTGCAACTGCAAAGCTGGAAAAGATTGACTT
ATTGTAATTGGGTAGAGGAGGTATTGAAACTGCATTGGAAGCAGCAGGG
AGAAATGGAATAGTGAATTGGTTGATGAATCGTCACAAACAACACTAC
CACTACTCTGTCAATCTTCAAACCTCAGAGAAATGAAGTTATGGTATC
25 TAAATTGTCGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTGAG
TTTCCAAACCTAACAAAGAGTCGATATATGGGGATGTGATAGGTTAGAAC
TGTATTACTAGTCCATGGTGGTAGTCTATTGCAACTCCAAGAGCTAC
GCAATGGAACTGCAGTCAGATAGAGGTCGTGATTGTCAGGATGCAGAT
GTTTGTGTTAGAAGAAGACAAAGAGAAAGAACATCTGATGGCAAGACGAATAA
30 GGAGATACTGTGTTACCTCGTCTAAAGTCCTGATATTAAAACACCTTC
CAWGTCTTAAGGGTTAGCTGGGAAGGAGGATTTCTATTCCCATT
TTGGATACYTTGAAATCTACRAATGCCAGCAATAACCACCTCACCAA
GGGAAATCCRCTACTCCACAGCTAAAAGAAATTGAAACAMATTGGCT
TCTTTATGTCAGGGAAAAAGACATCAACTCCTCTATTATAAGATC
35 AAACAACAGGTAAACCAAGATCTTGTGTTNNATAATTCTAAACNACA
TNTGAAAAGCTTCATGCAAGTTTTNGTTATATNGTCAAAACCGCAA
CCTACATTTCAGCTTANATTATGACTTTATGCAGGATTCAAACAA
GACTCTGATTAATGTGAAGTGAATATTAAAGGTAAATTATATTTCATGT
TCCTAGTNGCTATTAAATTAAAGGCCTTGTAGTCGNGATTGGATGT
40 ATTCTTCATGATGATGTCAATCTCTAAATACCCATTCAATTGTTGGT
AATGTTGACTCTATGTCAGGATGAATATTCAAGGGAAAGATTGTTCATCA
TATGAAGGACATTAAGAACATGGATGCTCTGAAGATGTTGGAACACA

RG2A deduced polypeptide sequence (SEQ ID NO:88)

MDVVNAIKPVVETLMVPVKHIGYLISCRQYMREMGIKMRGLNATRLGVEEHVN
 RNISNQLEVPAQVRGWFEEVGKINAKVENFPSDVGSCFNLKVRHGVGVKRASKIIEDI
 DSVMREHSIIWNDHSIPLGRIDSTKASTSIPSTDHHDEFQSREQTFTEALNALDPNHK
 5 SHMIALWGMGGVGKTTMMHRLKKVVKEKKMFNFIIAEAVVGEKTDPIAIQSAVADY
 LGIELNEKTKPARTEKLKWFVDNSGGKKILVILDDVWQFVLNDIGLSPLPNQGV
 DFKVLLTSRDKDVCTEMGAEVNSTFNVKMLIETEAQSLFHQFIEISDDVDPELHNIG
 VNI VRKGGLPIAIKTMACTLRGKS KDAWK NALL RLEHYDIENIVNGVFKMSYDNL
 QDEETKSTFLLCGMYPEDFDILTEELVRYGWLKLFFKVVYTIGEARTRLNTCIERLI
 10 HTNLLMEVDDVRCIKMHDLVRAFVLDMSKVEHASIVNHSNTLEWHADNMHDSC
 KRLSLTCKGMSKFPTDLKFPNLSILKLMHEDISLRFPKNFYEEEMEKLEVISYDKMKY
 PLLPSSPQCSVNLRVFHHLKCSLVMFDCSCIGNLSNLEVLSFADS AIDRLPSTIGKLK
 KLRLLDLTNCYGVRIDNGVLKKLVKLEELYMTVVDRGRKAISLTDDNCKEMAERS
 KDIYALELEFFENDAQPKNMSFEKLQRFQISVGRYLYGDSIKSRHSYENTLKVLEK
 15 GELLEARMNELFKKTEVLCLS VGDMD LEDIEVKSSSQLLQSSSFNNLRVLVVSCK
 AELKHFFTPGVANTLKKLEHLEVYKCDNMEELIRSRGSEETITFPKLKFLSLCGLP
 KLSGLCDNVKIIELPQLMELELDDIPGFTSIYPMKKFETFSSLKKEEV LIPKLEKLHVSS
 MWNLKEIWPCEFNMSEEVKFREIKVSNCDKLVNLFPHKPISLLHHLEELKVKNCGSI
 ESLFNIHLD CVGATGDEYNNSGVRIIKVISCDKL VNLFPHNPMSILHHLEELEVENC
 20 GSIESLFNIDLDCAGAIGQEDNSISLRNIKVENLGKLREVWRIKGGDNSRPLVHGFQS
 VESIRVTKCKKFRNVFTPTTTNFNLGALLEISIDDCGENRGND ESEESSHEQEIQIEILS
 EKETLQEATDSISNVVFPSCLMHSHFNLQKLILNRVKGVEVVFEIESESPTSREL VTT
 HHNQQQPIILPNLQELILWNMDNM SHVWKCSNW NKFFTLPKQQSESPFHNLTTIKI
 MYCKSIKYLFSP LMAELLSNLKH KIRECDGIGEV VSNRDDEDEEMTTFTSTHTTT
 25 LFPSLDSLTL SFL ENLKCIGGGAKDEGSNE ISFNN TTATTAVLDQFELSEAGGVSW
 SLCQYAREM RIEFCN ALSSVIPCYAAGQM QKLQVLT VSDCKGMKEV FETQLRRSSN
 KNNKSGAGEEGIPRVNNNVIMSLGLKILEISFCGGL EHIFTFSALES LRQLQELKITFC
 YGMKVIVKKEEDEYGEQ. TTTTTITKGASSSSSSSSKEVVVFPR LKSIELNDVPELV
 GFFLGKNEFR LPSLEEVTIKYCSKMMVFAAGGSTAPQLKYIHTELGRHALDQESGL
 30 NFHQTSFQS LYGD TLGP VTSEGTTCSFHNLIELYMEF NDAVKKIIPSSELLQLQKLEK
 IHVTYCNWVEEVFETA LEAAGRNGNNSGIGFDESSQTTTTLVNLPNLREMKLWYL
 NCLRYIWKS NQWTAFEF PNLT RVDI WGCDR LEHVFTSSMVGSLLQLQELR IWNCSQ
 IEVVT VQDADVC VEEDKEK ESDG KTNKEI LV PRLK S LILK HLP CLKG FSLG KEDFSF
 PLLDTLEIYKCPA ITFTKG NS TTPQLKEIETHFG FFYAGEKD INSSIIKIKQ QDFKQ
 35 DSD.CEVNIK

RG2B polynucleotide sequence (SEQ ID NO:89)

TTTTTTAAGATCAGGGATTCAAATTCAAGCCCTAGTGATTACAATTGTGTC
 TAAACTTCCCATA CCTTCACATTATTGTAAGTAACTTCTCAGTTCT
 40 CTCTTGGAAAGCTCCTTGGTATT TA ACTCGT GTTCTA AT ATTAACTCT
 GATAGTTATTTGCCAATCTACTATCTGCATGTCCGGTTATTGAATCCG
 AAGGC ACTGGAATCTTGGATTCCATTCCGTTGTGTTGGT GCAAAT

GAACGGAATTGAATTATGTAAGATTCCCTCAAAATCCATGTTAGGTATA
TCGTTGTTCTGGGATGGATGGTAAAGAACCGAATTCTCCTGTTCA
TTTAATGAAAGACCAAATTGACCTATAAACCTGTTAAAAAAATTACAT
TCCAGTTTCTTAACAAACTGAAAATGGTAAAGGAGTGTGATTGAATTCC
5 AATCTGTTCTGTCAAAACACGTGACGGAATTACAATTCTCAAA
TTTCATTTCTTAAATTGTTATTCCCTTCTTACAAAAACAAGGTAACG
AAACACCCGCTTACTTAATCATACTCCTACATGATGTAATGAAAAGGGT
ATAAAATGGTATTATTACAGGGATGAGTCACCATGGTCATGAAAGAAT
CATTAACCGCCCTTACCCAATTCATGTTGCCCTAAATATGATTAAA
10 GTAATATTGGCTTATGGGATTCAAGTTGACTTTTGTCGGCGAAGAAATA
ATGAAAATCTTCAATTCTAAAGTGTCTTACCACTGACATTCTAAGA
AAGAACCTGCTAGAAGAAGGTGGGTTGTTAGTCTTTACTCTTAAAT
GTGAAGACTGTTGAGTTATTATTATTGCAACTATGGACAACCTG
TTAGTTTTTTTCCCATTCCATTATGCGATTATTCTGA
15 AATAATTATCAAAACGCAGGAAACAATGTAGAATAACTGGTATAAT
TAATTATATAAAGTTATTAGGCTGAAATCTTGAGGCTACTATAATTAAAT
TATCATAATTGAAAATCATCAAATTGTATTCCATGTATATTATGTTAT
CAGATAATTAAATAATGTGAGCCACACAAATCCACATCATCAGACACCC
CACCTTATTGTCGGTACCTCACCACTGCATGATCCGACATCTCCA
20 ACCCCACCGACGACTTGGGGTCTCCTTAATATATCAATTATTCTGTAA
GTATTATTGTGAAATGTGTAATGTCATTACCTTTCTAATATA
TACAGAAACATAAATTAAATGAAATTCAACTGCGTTCATCTGCAT
TAAAAAAAAGACTGTACTGTTGCAATTATTACTTATAACCTGATTAA
TTAATTAAAGCGTAATTGATAATTGCATTAGGTTGTAATTGTGTT
25 TATAGGGAGGGTGGGGTCAACGGGAATCAAAGCACTATGAAAAGCAG
GGAAATACAAAAAATTACTCGAAACAAATTATTCAATTAAAGTGAGA
TAATAATGTTCTGATTAGATTATGAGAACTAGGGAGATTAAAGTGTATAT
CCCATTAAAAGAAATTGCATTATTAAATTGATCTTGATGACAA
AAATTAACTCGTGACAGGTTATATATCATACAAAATGAGTGGCTATGC
30 TTTCGTTCCAAAAGCAATTATAGTTATACTACACCTACAAATTAA
AAGGGGTTAAACATATCAAATACTTGATAAGTAATTATATAAATATGCA
TTAACCCCTCTAAAGAAAATGCTACTAAGCTGGACCATCTCAGAATTAC
AATCATAACCTTCCCTCAAAAAGATTGTTATATCATGTCATTGGC
ATTCAATTCTTTTCAACATTGATAGTTCTATTCTCAAAAATCGAGTT
35 CTCGTATTGTAAGGAAGATCAGAAGAGACTGTTCACACAGGTACTCT
TTTATTATTGATTCACATTGATATGTTATTGTTCTGCTTAATGG
TTTCGTCAGTCTAACTGCGCTGCTGATTAAATTCTTCACATTCTCC
ACGGATTTTAAATATTAGTTGTAATGAAACAAATTGGTAAGGAAAG
AAACATGGGAGTCTTCTAAAGTAAACCTAGATACTTAGGTTATAAGGG
40 TATATGCTAAAATGAACTATGCCATTACCTTGCCTTTCTTACTT
TTTATGTTTAAATGAAATCCAAGTTCTATGATCTGATGTGAGAAGA
ATAGGCATTAGAAAGGTAAAGGACGTACATAAAATTGATTAAATTAGTGA
TGTCTTGTATCATTATTACTCTCATAAAAAGCATATAGATCAAA

CACAAATTGCTACTTGTAGTGTAAACAACCTCGACTTAATAATGTTAATA
ATCAAGAGATTCTCTTGATTCAACTATTTCTAACCGAACAGCTCACTAA
AAACTCATATTGCTTGAGTCTGAGTGGTTATATTGGGTTTACATT
TAATTTTGTGCATGAATGTAAAATAGACTGCTTATTGATTCTTGTG
5 TTTCATTGAGTTGATTTCATTATTACTACCTTACAAATTGCTCAGTGAT
AGATTCCATTAAATTGCTAATTGGTTGCTTCTAAATATGTAGGAGCTA
CTAAAGCAAAATATCGAGCAATGTCGGACCCAACGGGATTGCTGGTG
CCATTATTAAACCAATTGCTCAGACGGCCTGGTCCCGTTACGGACCAT
GTAGGCTACATGATTCCCTGCAGAAAATATGTGAGGGTCATGCAGATGAA
10 AATGACAGAGTTGAATACCTCAAGAACATCAGTGAGAGAACACATTAGCC
GGAACACAAGAAATCATCTCAGATTCCATCTCAAACATAAGGAATGGTTG
GACCAAGTAGAAGGGATCAGAGCAAATGTGGAAAATTCCGATTGATGT
CATCACTTGTGTTAGTCTCAGGATCAGGCACAAGCTTGGACAGAAAGCCT
TCAAGATAACTGAGCAGATTGAAAGTCTAACGAGACAACCTCCCTGATC
15 AGTTGGACTGATGATCCAGTCCTCTAGGAAGAGTTGGTCCATGAATGC
ATCCACCTCTGCATCATTAAGTGTGATTCCATCAAGAGAGAAAATT
TTACACAAGCACTAAAAGCACTCGAACCCAACCAAAATTCCACATGGTA
GCC TTGTGTGGATGGGTGGAGTGGGAAGACTAGAATGATGCAAAGGCT
GAAGAAGGCTGCTGAAGAAAAGAAATTGTTAATTATATTGTTGGGGCAG
20 TTATAGGGAAAAGACGGACCCCTTGCCATTCAAGAACGCTATAGCAGAT
TACCTCGGTATACAACCAATCAATGAAAAACTAACGCCAGAACAGAGCTGATAA
GCTTCGTGAATGGTTCAAAAAGAATTCAAGATGGAGGTAAGACTAACGTTCC
TCATAGTACTTGACGATGTTGGCAATTAGTTGATCTTGAAGATATTGGG
TTAAGTCCTTTCAAATCAAGGTGTCGACTTCAAGGTCTTGTGACATC
25 ACGAGACTCACAAGTTGCACTATGATGGGGTTGAAGCTAATTCAATTAA
TTAACGTGGGCCTCTAACATGAAGCAGAACGCTCAAAGTCTGTTCCAACAA
TTTGTAGAAAACCTCTGAGCCCGAGCTCCAGAACAGATAGGAGAGGATATCGT
AAGGAAGTGTGCGGTCTACCTATTGCCATAAAACCATGGCATGTACTC
TTAGAAATAAAAGAAAGGATGCATGGAAGGGTCACTTCGCGCATAGAG
30 CACTATGACATTACAATGTTGCCCAAAGTCTTGAACGAGCTACCA
CAATCTCCAAGAACAGAGGAGACTAAATCCACTTTTAATGTGTGGTTGT
TCCCCGAAGACTTCGATATTCTACTGAGGAGTTGATGAGGTATGGATGG
GGCTTGAAGCTATTGATAGAGTTATACGATTAGAGAACGAAAGAACAG
GCTCAACACCTGCATTGAGCGACTGGTGCAGACAAATTGTTAATTGAAA
35 GTGATGATGTTGGGTGTCAAGATGCATGATCTGGTCCGTCTTGT
TTGGGTATGTTCTGAAGTCGAGCATGCTCTATTGTCAACCAGGTAA
TATGCCTGGGTGGCCTGATGAAAATGATATGATCGTGCACCTTGCAAAA
GAATTTCATTAACATGCAAGGGTATGATTGAGATTCCAGTAGACCTCAAG
TTCCCTAAACTAACGATTTGAAACTTATGCATGGAGATAAGTCGCTAAG
40 GTTCTCAAGACTTTATGAAGGAATGGAAAAGCTCCATGTTATATCAT
ACGATAAAATGAAGTACCCATTGCTCCTTGGCACCTCGATGCTCCACC
AACATTGGGTGCTTCATCTCACTGAATGTTCAATTAAAGATGTTGATTG
CTCTTCTATCGAAATCTATCGAATCTGGAAGTGCTGAGCTTGCAAATT

CTCACATTGAATGGTTACCTCCACAGTCAGAAATTAAAGAAGCTAAGG
TTACTTGATCTGAGATTGTGATGGTCTCGTATAGAACAGGGTGTCTT
GAAAAGTTTGTCAAACCTGAAGAATTATATTGGAGATGCATCTGGGT
TTATAGATGATAACTGCAATGAGATGGCAGAGCGTTCTAACACCTTCT
5 GCATTAGAATTGCGTTCTTAATAACAAGGCTGAAGTAAAAATATGTC
ATTGAGAATCTTGAAACGATTCAAGATCTCAGTGGATGCTCTTGATG
AAAATATCAATATGAGTAGGCCACTCATCGAAAACATGTTGCAATTGGTG
ACCAACAAAGGTGATGTATTAGACTCTAAACTTAATGGGTTATTTGAA
AACAGAGGTGCTTTTAAGTGTGCATGGCATGAATGATCTGAAGATG
10 TTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTCATTCTGCAATT
AAAGTTCTTATTATTCAAAGTGTAGAGTTGAGATACCTTTCAAAC
CAAATCTTGCAAACACCTTGCAAGACTGAGCATCTAGAAGTTGTGAAT
GTGAGAATATGGAAGAACTCATACATACTGGAATTGGGGTTGGAGAA
GAGACAATTACTTCCCTAAGCTGAAGTTTATCTTGAGTCAACTACC
15 GAAGTTATCAAGTTGTGCCATAATGTCAACATAATTGGCTACCACATC
TCGTAGACTTGATACTTAAGGGCATTCCAGGTTCACAGTCATTATCCG
CAGAACAAAGTGCACATCTAGTTGTTGAAGGAAGGGTAGATATG
TTCTTATGTTAATACAATTAAATAATTTCACCAAATTTCATAA
TATATCTGTAATTGATTGTATGATGTGTTATTGTTATATGTGGCTATT
20 AAGGGATGATTATTTCAGGGTGTGATTCTTAAGTTGGAGACACTCAA
ATTGATGACATGGAGAACTTAGAAGAAATATGGCCTGTGAACCTAGTGG
AGGTGAGAAAGTTAAGTTGAGAGCGATTAAAGTGAAGTAGCTGTGATAAGC
TTGTGAATCTATTCCCGCGCAATCCCATTGAGTCGTTATTCAACATTGA
GAGCTTACAGTCGAGAATTGGGTTCCATTGAGTCGTTATTCAACATTGA
25 CTTGGATTGTGCGGTGCAATTGGAGAAGAACACAAGAGCCTCTAA
GAAGCATCAACGTGGAGAATTAGGGAGCTAACAGAGAGGTGTGGAGGATA
AAAGGTGCAGATAACTCTCATCAACGGTTCAAGCTGTTGAAAG
CATAAAGATTGAAAAATGTAAGAGGTTAGAAATATTCACACCTATCA
CCGCCAATTCTATCTGGTGGCACTTTGGAGATTAGATAGAAGGTTGC
30 GGAGGAAATCACGAATCAGAACAGAGCAGGTAAACGCTTCAATTCACTTTC
TTAATTAAAGGACTAACGCTCCTGTTTGAATAATAAGAGGTGGG
ATGACTAAACTGGGCATCACAATTGCAACAAATGTTACAAACCATGAA
ACGTTCAAACCATTCTGAATTAAAGGTTCAATACAAGTCATTAAAAA
TATGGCTTAAATTCTATATTATGTATGTTACATGATTTCATTAG
35 AGATCATTATTATAATAGTAAGTTAAAGCAATTAAATCAGAACTAATT
CTAACTTAGCTAATAAATCGTTATAAATGTAATAATTACTTTAGTG
AAATAAGCAACGGATTAATAAGTTAACAACTTAAATGTCATTCTAAC
AAAAAAACTATTGGTCAGAAAACCGTAATTCAAGATAACTAAAATA
AAAATATTGACATTCACTAACAGAGCATTCTAAATGATTGCA
40 AATGAATAAAACTAAATTATACAGAAAATTCTTTATATATGTTATAC
AAAATTACAAATTGAAATTGGATATGTTAAATTACGGTTATAATTCTG
GTAATCACAAAGGGATATAATAAAAATTATTCTGTAGTCATTGTA
ATTGTACTAGTTATAACCCGTGGAACCATGAGTTCTAAAATTAGTTAA

ACTTTCATAATAAAAATTATAATTATTATTATTAAATAAAATTATTA
ATTAAGAGATATACAAAAATTAAAGTTATTATAACTCAAATTAAACA
TATAATTAGAAAATATGATCATAACTTCTGCAACTCTTCTTGTAT
TAAAATGACCAGAGAAGCTCTTAGTATATTCTAATCAAAGTCTCAAAMC
5 TAATGAAGCATATAATTGTGAAAATCAATTAGCATTAGGTTAACAGT
CACCAAATTCAAAGAATAATCCAATGCTTCATTACCACTATGGAGAAA
TATTTCTTAGTTAAATGAAATGAAAACAAACATTCAAACATAATTGTG
CTTATTAAACCAAGACCCATTACTTAGCCAAGAGTTAACAAAAAAA
TTACATTCATGTATCATTATTCACTGACTAGATATATGAACATGAAGGG
10 AGTTTTATAGAAAATATAATCATAGATATTCAACATAACTTCAGGGAAT
TCCTCAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCACCAAAG
AGAAGTTAGCCTAGCATGGCTAAACTCAAGAAACTAAAATAAGGATTAG
AAGTACCAAACATGTAGTAAGAATCACAGTAAAAGATGATGTTGTTCTG
ATGTTCTCTAAGTCTCAAGTCTCCAGTGCTCCTAATAATGCAAAGG
15 AGAGCCATTAAATTCTGTATGTATTGATCCCTCAAAAGCTGCACCAACCT
CCCTTAAATAACACTCAAAGCAAAATGACAAAATTGCCCTGAAGGACC
CTATGTGGGTGCCTGCGCGGGTGGAGCTGCATACGAAAGGTCTTGGTC
TTTGTGAGGGTGTAGTTGCGGGATAGCTTGTGCGATGCTCCGCGCG
TTCACGCACATGTGCACAGGTGATGCATGGTGTGCGTTCTTGAGTTT
20 GAGCCTCCGATGCTTAGTCCACTTGGCCAATTGAGTCCAATCAGCTTA
TAACCCATTTCCTCAAGTTATCTCAAGTTAACGCCAATTGCCCTCT
CCA-AATCATCCATAACTTCACAGAACATGCCCTCATCTTAATCCGGAT
GCACAATTATTCTCCCGTCTCATTTAACAGCAAGATAACCACCTTCTCAT
GCTTCATCCATCAATAGTACACTTCATGTATCTACTAGTTATTAA
25 GTCCACAACTCTGTTGCTCCAAATTAAATTATCTCATTAGTCCCG
TTCCGCTAGTTCCCTAAATTGCAATTAAAGCTCAGAGAAAATTAAGT
ACCCGAAATGGTCATAAAATAACAAAAAGGAAAATATGCATGAAGATTAA
CTA-AATGATGAACGAAATATGCTAAATAGACTATAAAATGAAGTAAATA
AAATGAAATTATCGCACTCCGACCACCCCTATAGGCTGTAGTCCACCCA
30 CCCTCATTCTGTACCAATATGGGATGGAAACATCATTAATTAGCCA
AAAAGCTAACATATAAGGGTTAGTGCACAAAGGTAAGTACTAAAGATGAA
ATAATCCATTTCCTGTATATACACAAACACACATAGGGCAGACGT
AGGATTCAAAAGTACAGATTGTTGGCAGCATAGTGTGCTGGTGACA
TTTTTTTTTTTACGTAGTGGCACACAGTAGGAAAACGAAAAT
35 TCGAAATTTCACAATTGTCTAAAAAACAGTGGTTGTTGGTGCAC
TATGGACACCAAAGTTGAACTGCCCCCACGCGCGCACACACACACACA
CATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAAAGAAAAGAGAGA
GAGAGTTGGATGTGATACTTCTTTAGGAAAATGGAGTTATCTTG
ATATTGTATTTCATAATGTAATTATATTAAATCATTTAGTTATA
40 AGTTTATTATTTGATATGAAAAAAAGCTTTATACATTGGATT
AACATAAAAATCCAACAATTAAATCAAAAGACCAMACATGTGGACAMW
TATGTATATAAWTAATTACAATAGTCTTAGGAATAGNATTATATAT
AATTAATTCTCAATGGTCTTAGGAATAGTAAGTTCTTATATTCAAACCTT

TNGCCACAATTCTTGKTTACTTWGACACTTYCCTCTCTAATTATATA
TATATATATATATATATATACACACACACACACACTAG
ATGTGTGCCCGCGCAAAGCAGTGACGTNNNGAGAANACTTCTTAAGCA
TAAATAATTATTATTTTATTGGGTATTATATAATAAAAAATTACAA
5 CTTTAAATAAAATATTATGTTATACTTATTTATATTGCTTGTAT
ACTATTAATATAATAATTAAATTTATGTCTAATTATGAAATGTAAT
TAATTAAATACATGAATTAAATTTTAAATTTCAAGTTGCTCAA
ATTGAGTTCTTAATTATTTTTAATTCAAGTATTCAAAACTTTGGTA
AGTATTAAAGAATTATTATGCATAATTGATTATACAAAAACTTGTA
10 ACTTATACATCTTAAATTCAAGATATAACTAACATGTTACAATATAT
AT
TAAAGCGCAAAGGTCAAGGAATAGAAATATTCTATTATTCTACGTTT
GCCACAAAAGTTGAACACTTGCCACTTTGTCCCTCTAACCTTT
CAATGTTTGCACAAAAGTCCAAAACATTGCCACTTGATCATTCCTC
15 AACTTTCACCGCAATTAGTTGTGGAGTTGGCAGTTGATCCCCCTAA
CTTCGATATTCTCTACTGCTAGCCAAAAGGGTCCAGAGTTCACACTT
TTGGTCCCTGACAGTAACCAAATGTGAGATGTCAAATTGGCCACATTA
GTTTGTGGAGTTGCCCTTGGTCCCCCACATTGATATTCTANTATA
CGACCTTATTTNTCAAATAACACACGTATATTAAATTACCAATTATA
20 GAAATAGATATCAAATAAAAGTATTGTAACACTGTGTAAGAACGGTGCTA
CTATAGGTAAAATAACATTCAAAGTACGATATCCTAATTGGAAAAAG
AGTTTAAAAAAATAACGACTAGGGCGAGTTTTTACAAGTTGTAT
CAAATCATATCAAATTAAAGGTGGAACGGTGACCACATTAACAGAAAT
GTAATTATTCTTGATTGATAATTAAATATTGTTGTGATCTAT
25 GTATTAAAAGTAAACAACAAAGAACATAATCCAAAACCTAAATTGCAA
GTCTGCCAATTCTCTACTGCTACTAGCCTCACTTACGATGGCGTTACGT
CGCTCTCACTGCTACAACCCTTGTGCTACTCATTACAATAACGAA
AAGTTGAATATCCATATTGATGTGGAATTGAACGAATCTCGT
CAAATTGATTGATGGATTGAGTAGAAGTTGGCAGAACGG
30 GAAATGATGGTCTGCAAGTGGTATAAAACTGATTCTGAGTTATTACTATA
TATGTAGCCTCTTACAACGACCAAGGTTCTCCAGGTACCATTTGATC
TTTTAGAACCTAGTTCTGAAACACCCCTGATTGGATCAAATATCACC
AACAACTCTAAAAACTTGATTAATCAATTGTTCTCATCTTGATAAC
AAGTGGATGATTCTACTTAGATTAACCTGAAAAAAAAGGTCCATGTG
35 CGCTGGTGGATCTGGTAAATGAAGATGGAAGGGAGAGCTGACTTAAAG
ACACAAACACGTCAACATCTTATTAAATTGCTTTGGT
GTATTCTTTCTTCTATTCTTCTTGATCTCCAGATGGTATGT
GGTGTGGATAATTACACCTAGAGATTGGAACGATGGAAGGGTCTGT
GATTATGGCTGGCGAGTTTACTTATTAAACTCAATTCAACCTAAATT
40 CTGATTCTGTTGAAAATAAGTTGCATCTTATTGTTGATTATCTTGT
TGCATAGGATCCTTAGCATTTAATAATTATTGAAGGTGAAAGATC
CAACTATTAGCTGTTGCATTTCATCATTTGCAACTGTTCTG
AAAAAAAAACCTAAAATAACCATTCAAAATTATA

AGAGAGAATTGTAATGGACATGGAATCATAAACATTAACACAGTCAG
TAAACAAGTTGCTAATTACATTCTGTGCTGTGCAGATTGAAATTCTATCA
GAGAAAGAGACATTACAAGAAGCCACTGGCAGTATTCAAATCTGTATT
CCCATCCTGTCTCATGCACTCTTCATAACCTCCGTGTGCTTACATTGG
5 ATAATTATGAAGGAGTGGAGGTGGTATTGAGATAGAGAGTGAGAGTCCA
ACATGTAGAGAATTGGAACAACACTCGCAATAACCAACAACAGCCTATTAT
ACTTCCCTACCTCCAGGATTGTATCTAAGGAATATGGACAACACGAGTC
ATGTGTGGAAGTGCAGCAACTGGAATAAATTCTCACTCTCCAAAACAA
CAATCAGAATCCCCATTCCACAACCTCACAACCATAAAATTCTAAATG
10 CAAAAGCATTAAAGTACTTGTTCGCCTCTCATGGCAGAACTTCTTCCA
ACCTAAAGGATATCCGGATAAGTGAGTGTGATGGTATTAAAGAAGTTGTT
TCAAACAGAGATGATGAGGATGAAGAAATGACTACATTACATCTACCCA
CACAAACCACCACTTGTCCCTAGTCTGATTCTCTCACTCTAAGTTCC
TGGAGAATCTGAAGTGTATTGGTGGAGGTGGTCCAAGGATGAGGGGAGC
15 AATGAAATATCTTCATAAAACCACTGCAACTACTGCTGTTCTGATCA
ATTGAGGTATGCTTGTACATATTCAATTATTATTAAATTCTTTT
TATTGCAATATTCTATAAAATAACATTACACCCACTATACTAAGAT
AAATAATTACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTCAGA
AACTCTAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCT
20 TTTGATGGTAATATAGGCAATTAAAGTTTATTCTGTTAAAGCAGTAT
TTAGCAAGTACTGCCAGTAGGAGAGGAGAATATCACCTTTGTGAAAAT
CTGGTCATTGTACCCAGAATTAGTTAAATGTAACATTAGATATTAGG
GGACATCAGGTGACAGATATTGAGAATAGAACAAATATAATTACCC
AAAACATTCTAAGGTTATTCTGTTAAATATGTGCTTCTTGATT
25 CATTGAATTGCAATTCTATATTAGGTGGTAAAGTGATTGTCTCTCA
ATAAAATCCGAAATTAAATTAAAAAGAAAAAAACAAAAGTAATTCTTG
TATGGAGAGCACTGGTATCATTTAGTATATAAAAAAACTAGATTGAAAT
TAAGTTCTTATATAAAAGCTGTATATAGTTAATTAGTTACATCA
TTTTCCATGTGGTGTGCAGTTCTGAAGCAGGTGGTGTGAGGAG
30 TTTATGCCAATACGCTAGAGAGATARAAATAGKGGATGCTATGCATTGT
CAAGTGTGATTCCATGTTATGCAGCAGGACAAATGCAAAAGCTCAAGTG
CTGAGAATAGAGTCTGTGATGGCATGAAGGAGGTATTGAAACTCAATT
AGGGACGAGCAGCAACAAAAACACGAGAAGAGTGGTGCAGGAAGGAA
TTCCAAGAGTAAATAACAATGTTATTATGCTCCCAATCTAAAGATATTA
35 AGTATTGAAATTGTGGGGTTGGAACATATATTACACATTCTCTGCAC
TGAAAGCCTGAGACAGCTCAAGAGTTAAAGATAAAATTGCTACGGAA
TGAAAGTGTGAGAAGAAGAAGATGAATATGGAGAGCAGCAAACA
ACAACAAACAACGAAGGGGGCATCTTCTTCTTCTTCTTC
TTCTTCTCTAAGAAGGTTGTGGCTTCCCTGTCTAAAGTCCATTGTAT
40 TGGTCAATCTACCAAGAGCTGGTAGGATTCTTCTTGGGATGAATGAGTTC
CGGTTGCCTTCATTAGATAAACTTAAGATCAAGAAATGCCAAAATGAT
GGTGTACAGCTGGGGTCCACAGCTCCCCACTCAAGTATATACACA
CAAGATTAGGCAAACATACTCTGATCAAGAATCTGGCCTTAACTTCAT

CAGGTATATATATTTCTTAATTGGCATCATCTAATTAAGAAAGATAT
CATTCCGCCAAGTAAATTACTTCAAACACATTACACTGGTTCAGTC
TAAGTTATGTTCTAAGAAGGCCAAATGGGAAAGCAAGATAGGGAA
AAATAGTGATTCAGTGGAAAGGGTATTTAGGCATTTCTGTCAAAAG
5 TTGTTATTGCAGGCCCTTACTGACCTGGAAATCGTGTGGGAGGAGCATT
ATTATTCTGATTGCTTCTTATCATTTCCTTAGCCTCTCGAAC
AGCTAGAAACCCTTTAATCTTTGATTTCATGACGAAATTTCCT
GTTACTCCATTGATTGTTCTCATGGTTCTAAGTGAGTTATTGGCT
CATCTGTTACTTCTTGTATTGTTATTCATATCATGTTGCCTTGAA
10 TCAAGCTTTCCATTTCACCAGGGCAAAAGGTCAAAAGTAACCTACTT
TATGAGATCAAAAACAGCAACCCATCGGATAACTTTAGTTGGAGTTAAT
AGTTACAATTACCATTGTGATTAATAATTATAATCTTGATTAATTCA
AAAAAATTGGTACAGCACATATGACATTCAAAAGGTTTGTTGACA
TATATATGCCTCTGGCGTTTCTTATTGGACTTGCAGACCTCATTCAA
15 AGTTTATACGGTACACCTGGGCCCTGCTACTCAGAAGGGACAACCTG
GTCTTTCTATACTTATCGAATTAGATGTGGAAGGTAATCATGATGTTA
AAAAGATTATTCCATCCAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAG
ATTAAATGTAAGGTGGTGTAAAAGGGTAGAGGGAGGTATTGAAACTGCATT
GGAAGCAGCAGGGAGAAATGGAATAGTGGATTGGTTGATGAATCGT
20 CACAAACAACCACTACTCTGTCAATCTTCAAACCTTAGAGAAATG
AACTTATGGGTCTAGATTGCTGAGGTATATGGAAGAGCAATCAGTG
GACAGCATTGAGTTCCAACCTAACAGAGTTGATATCTATAATGTA
AAAGGTTAGAACATGTATTACTAGTCCATGGTTGGTAGTCTATCGCAA
CTCCAAGAGCTACATATATCCAAC TGCA GTGAGATGGAGGAGGTATTGT
25 TAAGGATGCAGATGATTCTGTAGAAGAAGACAAGAGAAATCTGATG
GGGAGACGAATAAGGAGATACTTGTGTTACCTCGTCTAAACTCCTTGATA
TTAAGAGAACTTCCATGTCTAAGGGTTAGCTGGGAAGGAGGATT
TTCATTCCCATTATTGGATACTTTAAGAATTGAGGAATGCCAGCAATAA
CCACCTTCACCAAGGGAAATTCCGCTACTCCACAGCTAAAAGAAATTGAA
30 ACACATTGGCTCGTTTGTGCTGCAGGGAAAAAGACATCAACTCTCT
TATAAGATCAAACAACAGGTAAATCAGATCTTGTGCTTAAATATTCA
TTAAACTACATTGAAAAGCTTGTGCAAGTTTTTGTATATTGTCA
AAAACCGAACCTACATTTCAGCTTATATTATGTACTTTATGCAGGA
GTTCAAACAAGACTCTGATTAATGTGAAGTAAACTAAAGGTAAATTAT
35 ATTTCATGTTCTAGTTGCCTATTAATTAATTGCCTTAGTTGATGAT
TTTGGATGCATTCTCATGATGATGTCATCTTCTAATACCCATTGAT
TGTTGGTTGAATGTTGACTCTATGTCTGATGAATATTCAAGGGAAAGAA
TTGTTCATCATGAAGGACATTAAGAAGAACATGGATGCTATGAAGAT
GTGGGAAAACAA

RG2B deduced polypeptide sequence (SEQ ID NO:90)

MSDPTGIAGAIINPIAQATLVPVTDHVGYMISCRKYVRVMQMCKMTELNTSRISVEE
 HISRNTRNHLQIPSQTKEWLDQVEGIRANVENFPIDVITCCSLRIRHKLGQKAFKITE
 QIESLTRQLSLISWTDDPVPLGRVGSMMASTSASLSDDFPSREKTFTQALKALEPNQK
 5 FHMVALCGMGGVGKTRMMQRLKKAEEKKLFNYIVGAVIDGEKTDFAIQEAIADY
 LGIQLNEKTKPARADKLREWFKKNSDGGKTFLIVLDDVWQLVDLEDIGLSPFPNQ
 GVDFKVLLTSRDSQVCTMMGVEANSIINVGLTEAEAQSFLQQFVETSEPELQKIGE
 DIVRKCCGLPIAKTMACTRNRKRDAWKDALSRIEHYDIHNVAPKFETSYHNLQ
 EEETKSTFLMCGLFPEDFDIPITEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQ
 10 TNLLIESDDVGCVKMHDLVRAFVLGMFSEVEHASIVNHGNMPGPDENDMIVHSC
 KRISLTCKGMIEIPVDLKFPKLTILKLMHGDKSLRFPQDFYEGMEKLHVSYDKMKY
 PLLPLAPRCSTNIRVLHLTECSLKMFDCCSIGNLSNLEVLSFANSHIEWLPSTVRNLK
 KLRLLDLRFCDGLRIEQGVLKSFVKLEEFYIGDASGFIDDNCNEMAERSYNLSALEF
 AFFNNKAEVKNMSFENLERFKISVGCSFDENINMSSHSYENMLQLVTNKGDVLDSK
 15 LNGFLKTEVLFLSVHGMNDLEDVEVKSTHPTQSSFCNLKVLIISKCVELRYLFKL
 NLANTLSRLEHLEVCECENMEELIHTGIGGCGEETITFPKLKFLSQLPKLSSLCHN
 VNIIGLPHLVDLILKGIPGFTVITYPQNKLRTSSLKEGVVIPKLETLQIDDMENLEEIW
 PCELSGGEVKLRAIKVSSCDKLVNLFPRNPMSSLHHLEELTVENCGSIESLFNIDLD
 CVGAIGEEDNKSLLRSINVENLGKLREVWRIKGADNSHLINGFQAIVESIKIEKCKRFR
 20 NIFTPITANFYLVALLEIQIECGGNHESEEQIEILSEKETLQEATGSISNLVFPSCLMH
 SFHNLRVLTLDNYEGVEVVFEIESESPTCRELVTRNNQQQPIILPYLQDLYLRNMD
 NTSHVWKCSWNKFFTLPKQQSESPFHNLTTINILKCKSIKYLFSPLMAELLSNLKDI
 RISECDGIKEVVSNRDEDEEMTTFTSTHTTLFPSLDSLTSFLENLKCIGGGGAK
 DEGSNEISFNNTTATTAVLDQFELSEAGGVWSLCQYAREIEIVGCYALSSVIPCYAA
 25 GQMQL

RG2C polynucleotide sequence (SEQ ID NO:91)

ATAATATTACACAAAGGTAAACGTCTTAATTAAATTACGATAACGAGACAGA
 CTTTTCACTCGGACATNAACGGTCTATTCTAACCTNNANTAAATTNAAT
 30 GAATTAGGATGTGCTAATATGCATGTAANATTGCTACCGTCATCTTC
 AAATGACCATATTTTATGTATTATAATGAATCAATGAAAAACCGGATT
 TCTATTAAAATTCTAAACTTCATCTTAAGCCAGGGTGAATACAAT
 TGCTAGATCCACTGTTAATTCCATCGAATTATGCCTGATCAATTGTTGG
 CTGCCTACGATGCAGGTGCTACCACAAGAATATGCCATGGAAACTGCTA
 35 ATGAAATTATAAACAGTTGTCAGTTCTCATGGTCTCTATTAAACGAT
 TACCTACGCTACCTCGTTCTGCAGAAAGTACATCAGTGACATGGATT
 GAAAATGAAGGAATTAAAAGAAGCAAAGACAATGTTGAAGAGCACAAGA
 ATCATAAACATTAGTAATCGCTTGAGGTTCCAGCAGCTCAAGTCCAGAGC
 TGGTTGGAAGATGTAGAAAAGATCAATGCAAAAGTGGAAACTGTTCTAA
 40 AGATGTCGGCTGTTGCTTCAATCTAAAGATTAGGTACAGGGCCGGAAAGGG
 ATGCCTTCAATATAATTGAGGGAGATCGACAGTGTCAATGAGACGACACTCT
 CTGATCACTGGACCGATCATCCCATTCTGGGAAGAGTTGATTCCGT

GATGGCATCCACCTCTACGCTTCAACTGAACACAAATGACTCCAGTCAA
GAGAGGTAAGGTTAGTGAAGCACTCAAAGCACTGAGGCCAACACATG
ATAGCCTTATGTGGAATGGGGGAGTGGGAAGACCCACATGATGCAAAG
GCTGAAGAAGGTTGCCAAGAAAAGAGGAAGTTGGTTATATCATCGAGG
5 CGGTATAGGGAAATATCGGACCCCATTGCTATTCAAGCAAGTTGTAGCA
GATTACCTATGCATAGAACTGAAAGAAAGCGATAAGAAAACAAGAGCTGA
GAAGCTTCGTCAAGGGTTCAAGGCCAATCAGATGGAGGTAACACTAAGT
TCCTCATAAATATTGGATGATGTCAGTCCGTGATCTAGAAGATATT
GGTTAAGCCCTCTCCAATCAAGGTGTCAGTCAAGGTCTTGTGAC
10 TTCACGAGACGAACATGTTGCTCAGTGTGGGGTTGAAGCTAATTCAA
TTATTAAACGTGGGACTTCTAATTGAAGCAGAACAGCTTGTCCAG
CAATTGTAGAAACTCTGAGCCGAGCTCCACAAGATAGGAGAAGATAT
TGTTAGGAGGTGTTGCGGTCTACCCATTGCCATCAAAACCATGGCGTGT
CTCTAAGAAAATAAAAGAAAGGATGCATGGAAGGATGCACTTCTCGTTA
15 CAACACCATGACATTGTAATGTTGCTACTGCAGTTTAGAACCAAGCTA
TGAGAATCTCCGGACAAGGAGACAAAATCTGTTTTGATGTGTGGTT
TGTTCCCGAAGACTTCAATATTCTACCGAGGAGTTGATGAGGTATGGA
TGGGGCTTAAAGTTATTGATAGAGTTATACAATTATAGAACAGAAA
CAGGCTCAACACACTGCATTGACCGACTGGTGCAGACAAATTACTAATTG
20 GAAGTGATAATGGGTACATGTCAAGATGCATGATCTGGTCCGTGCTTT
GTTTGGGTATGTATTCTGAAGTCGAGCAAGCTCAATTGTCAACCATTG
TAATATGCCCTGGGTGGCCTGATGAAAATGATATGATCGTCACTTGC
AAAGAACATTCTAAACATGCAAGGGATGATTGAGTTCCAGTAGACCTC
AAGTTCTAAACTAACGATTGAAACTTATGCATGGAGATAAACGCT
25 AAAGTTCTCAAGAACATTATGAAGGAATGAAAAGCTCCGGTTATAT
CATACCATAAAATGAAGTACCCATTGCTTCCCTTGGCACCTCAATGCTCC
ACCAACATTGGGTGCTTCATCTCACGGAATGTCATTAAAGATGTTGA
TTGCTCGTGTATTGAAATCTATGAACTGGAGTGTGAGCTTGCTA
ATTCTGCATTGAGTGGTACCTCACGGTCAGAAATTAAAAAGCTA
30 AGGTTACTTGATTGAGATTGTTATGGTCTCCGTATAGAACAGGGTGT
CTTGAAAAGTTGGTCAAACCTGAAGAACATTATGGAAATGCATATG
GGTTTATAGATGATAACTGCAAGGAGATGGCAGAGCGTCTTACAACCTT
TCTGCATTAGAATTGCGTTCTTAATAACAAGGCTGAAGTAAAAATAT
GTCATTGAGAACATTGAAACGATTAAAGATCTCAGTGGATGCTTTG
35 ATGGAAATATCAATATGAGTAGCCACTCATAACGAAAACATGTTGCGATTG
GTGACCAACAAAGGTGATGTATTAGACTCTAAACTTAATGGGTATTTT
GAAAACAGAGGTGCTTTTAAGTGTGCATGGCATGAATGATCTTGAAG
ATGTTGAGGTGAAGTCGACACATCCTACTCAGTCCCTTCATTCTGCAAT
TTAAAAGTCCTATTATTCAGGTTGAGATACCTTTCAA
40 ACTCAATGTTGCAAACACTTGTCAAGACTTGAGCATCTAGAAGTTGTA
AATGCAAGAACATGGAAGAACTCATACATACTGGGATTGGGGTTGTGGA
GAAGAGACAATTACTTCCCCAAGCTGAAGTTTATTTGAGTCAACT
ACCGAAGTTATCAGGTTGCCATAATGTCAACATAATTGGGCTACCAC

ATCTCGTAGACTTGAAACTTAAGGGCATTCCAGGTTCACAGTCATTTAT
CCGCAGAACAGTTCGAACATCTAGTTGTTGAAGGAAGAGGTAGATAT
ATGTTCTTATGTTAATACAATTAAACAATATTTCAACCAATTTC
TAATATATCTGTAATTGATTGTATGATGTGTTATTGTTATATGTGGCT
5 ATTAAGGGATGATAATTTCAGGTTGTGATTCTAACAGTGGAGACACTT
CAAATTGATGACATGGAGAACTTAGAAGAAATATGCCCTGTGAACCTAG
TGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAAGTGAGTAGCTGTGATA
AGCTTGTGAATCTATTCCGCGCAATCCCATGCTCTGTCATCATCTT
GAAGAGCTTACAGTCGAGAATTGCGGTTCCATTGAGTCGTTATTCAACAT
10 TGACTTGGATTGTGCGGTGCAATTGGAGAAGAAGACAACAAAGAGCCTCT
TAAGAAGCATCAACGTGGAGAATTAGGGAAAGCTAACAGAGAGGTGAGGAG
ATAAAAGGTGCAGATAACTCTCATCTCATCAATGGTTCAAGCTGTTGA
AAGCATAAAGATTGAAAAATGTAAGAGGTTAGAAATATATTACACACCTA
TCACCGCCAATTATCTGGTGGCACTTTGGAGATTAGATAGAAGGT
15 TGGAGGAAATCACGAATCAGAACAGAGCAGGTAAACGCTTCAATTCACT
TTCTTAATTAAATTANGACTAACGCTCCTGTTTGAATAATAAAGAGGT
GGGATGACTAAACTGGGCATCACATTGCAACAAAATGTTACAAACCAT
GAAACGCTCAAACCATTCTGAATTAGGTTCAATACAAGTCATTAA
AAATATGGCTTAAATTATATTATATTATGTATCAACATGATTTCATT
20 AGAGATCATTATTATAATAGTAAGTTAAAGCAATTAAATTAGAACTAA
TTCTAACTTAGCTAATAAAATCGTTATAAATGTAACATAATTACTTTAG
TGAAATAAGCAACGGATTTAATAAGTTAACAAACTTAAATGTCATTCTA
ACA AAAA ACTATTGGTCAGAAAAACTGTAATTCAAGATAACTAAA
TAAAATATTGACATTCACTAACAGAGCATTCTAAATATGATTGCA
25 AATGAATAAAACTTAAATTACAGAAAAGATTATATATGTTATAC
AAAATTACAAATTGAAATTGGATATGTTAATTACGGTTATAATTCTG
GTATCACAAAGGGATATATAATAAAAATTATTTCTGTAGTCATTAT
AATTGTTACTAGTTATAACCGTGGAACCATGAGTTCTAAAATTAGTTA
AACTTCTATAATAAAAATTATAATTATTATTATTTAAATAAATTATT
30 AATTAAAGAGATATCAAAATTAAAGTTATTATAACTTCAAATTAAAC
ATATAATTAAAAATATGATCATAACTTCCGCAACTCTTCTTTGTA
TTAAAATGACCAGAGAAGAGCTCTTAGTATATTCTAAATCAAAGTCACAA
AACTAATGAAGCATATAATTGTGAAAATCAATTAGCATTAGGTTAA
GAGTCACCAAATTCAAAGAGTAATCCAATGCTTCACTACCATGGAG
35 AAAATATTCTTAGTTAAATGAAATGAAAACAAACATTCAAACAAATT
GTTGCTTAAACCAAAGACCCATTACTTAGCCAAGAGTTAACCAAAA
AAAATTACATTGATCTATTATAATGACTAGATATATGAATATGA
AGGGAGTTTATAGAAAATATAATCATAGATATTCAACATAACTTCATG
GAATTCCCTCAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTCACCC
40 AAAGAGAAGTTAGCCTAGCATGGCTAAACTCAAGAAAATAAGGA
TTAGAAGTACCAAACATGTAGTAAGAATCACAGTAAAGATGATGTTGTT
CTTGATGTTCTCAAGTTCTCAAGTCTCCAGTGTGCTCTAATAATGCA
AAGGAGAGCCATTAAATTGATGATCCCTCAAAAGCTGCACCA

ACCTCCCTAAATAACACTCAAAGCAAAATGACAAATTGCCCTGAAG
GACCCTATCGGGTGCCTGCGCGGTGGAGCTGAATATGAAAGGTCTT
GGTCTTGAGGGTGTGCGGGTAGCTTGTGCATGCTCCGC
5 GCGGTTCGGCACATGTGACAGGTGATGCATGGTGTACGTTCTGAC
TTTGAGCCTCCGATGCTTAGTCCACTGGCCAATTGAGTCAAATCAA
CTTATGACCCATTTCCTCAAGTTATCTCAAGTTAACGCCAATTGCC
TTCTCAAATCATCCATAACTCACAGAACGCCGTCATCTTAATCCC
GAATGAACAATTATTCTCCGCTTCATTAAAGCAAGATAACCACCTTCT
10 TCATGCTTCATCCATCAATAGTACACTCATGTATCATCTACTAGTTA
TTTATGCCACAGTCCTGTTGCTCTCAAATTAAATTATCTCATTAGTT
CCCGTTCCGCTAGTTCTTAAATTGCAATTAAAGCTCACAGAAATATT
AAAGTACCGAAATGGTCATAAAATAACAGAAAGGAAATATGCATGAAGA
TTAACTAAATGATGAACGAAATATGCTAAAATAGACTATAAAATGAAGTA
AATAAAATGAAATTATCGCACTCCGACCACCTTATAGGCTTAGTCCA
15 CCCACCCCTCATTCTGTACCAATATGGATGGAACATCATTAATTAA
GCCAAAAAAACTAACATATAAGGGTGAGTGACAAAGGTAAAGTAAAGA
TGAAAAAAATCCATTTCCTGTATATACACAACACACATAGGGCAG
ACGTAGGATTCATAGTACAGATTGTTGGCACATAAGTGTGCTAGT
GACATTTTTTTCTTACGTAGTGGCACAAACAGTARAACRAAA
20 AATTGAAATTTTTACAATGTGCCTAAAAAACAGTGGTGTGGTGC
CACTATGGACACCAAAGTGAAC TGCCCCCTGCGCGCACACACACAC
ACATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGTTG
GGATGTGATACTCTTGGAAAATGGAGTTATCTTGATATTGTAT
TTTTTAATGTAATTATATTTAATCATTAGTTATAAGTTTATT
25 TATTTKGATATGAAAAAAAGTCTTTATACATTGATTAAACATAAAA
ATCCAACAATATTAATCAAAAGACCAACATGTGGACAATTATGTATAT
AATTAATTCAAAATAGTCTTAGGAATAGNATTATATATAATTAAATTCA
TCAATGGCTTAGGAATAGTAAGTTCTATATTCAAACNTTGCCACAN
TTCTTGNTTACTTNGACACTTYCTCTMWNNANWMWWTWATATATAT
30 ATATATATATATAHAHAHAHAVACACACACTAGATGTGTGCCMGC
AAGCAGTGACGTNNNGAGAANACTTCTTAAGCATAAAATTATTATA
TTTTTATTGGTATTATATAATAAAAAATTACAACCTTTAAATAAAATA
TTTATGTTACTTATATTATATTGCTGTACTATTAAATATAATA
AATTAATATTATGTCTAATTATGAAATGTAATTAAATTAAACATG
35 AATTAATATTAAAATTTCAGTTGCTCAAATTGAGTTCTTAAT
TATTGACCAAACATGTGGACAATTATGTATATAATTAAATTACAATAGTC
TTTAGGAATAGTATTATATATAATTAAATTCTCAATGGCTTAGGAATA
GTAAGTTCTTATATTCAAACCTTGCCACAATTCTTGCTTACTTGAC
ACTTTCTCCTAACTTACATATATAATTAAAGCGCAAAGGTC
40 ATAGGAATATAATATTCTATTATTCTACGTTGCCACAAAAGTTGA
ACACTTGCCACTTTGTCCCTCCTAACCTTCAATGTTGCGACA
AAAGTTCCAAAACTTGCCACTTGATCATTCTCAACTTTCACCGCAT
TAGTTGTGGAGTTGGCAGTTGGCCCTCTAACCTCGATATTCTCTAC

TGCTAGCCAAAAGGGTCCAGAGTTCACACTTTGGTCCCTGACAGTA
ACCAAATGTGAGATGTCAAATTTGCCACATTAGTTGTGGAGTTGTCC
CTTTGGTCCCCCACATTGATATTCACTATACGATCTATTTCCTC
AAATAACAACACGTATATTAACTAATGATAGAAATAGATATCAAAT
5 AAAGTATTGTAAACACTGTGTAGAGTTTTTACAAGTTGTATCAA
TCATATCAAAATTAAAGGTGGAACGGTGUACCACATTAACCAGAAATGTAA
TTTATTCTTGATTTGATAATTAAATTTGTTGTGATCTATGTAT
TTAAAAGTAAACAACAAAGAACATAATCCAAAACCCTAAATTGCAAGTCT
CGCCCATTCTCATCACTAGCCTCACTTACGATGGCGTACGTCGCT
10 CTCTCACTGCTTACAACCCTTGTGCTACTCATTACAATAACGAAAAGT
TGAATATCCATATATTATTGGATGTGGAATTGAAACGAATCTCGTCAA
TTTTGATTTAGTTGATGGATTGAGTAGAAGTTGGCAGAACGGGAAT
GATGGTCTGCAAGTGGTTAAACTGATTCTGAGTTACTATATATG
TAGCCTCTTACAACGACCAAGGTTCTCCAGGTACCATTGATCTTT
15 TAGAACTTAGTTCTGAAACACCCTGATTGGATCAAATATCACCAACA
ACTCTAAAAACTTGATTAATCAATTGTTACTCATCTTGATAACAAGT
GGAATGATTTCTACTTGAAAAAAAAGGTCCATGTGCGTCTGGTGGATCT
GGTAAATGAAGATGGAAGGGAGAGCTGACTTAAAGACACAAACACGTCA
CCATATCTTTATTAAATTTCCTTTCTTCTTCTTCTTCTTCTTCTT
20 CTTGATCTCCAGATGGTATGTGGTGGATAATTACACATAGAGATTGG
GAACGACTGTGATTAGAGAGGACGTGGCTGGGTTGAGGATGGTTAT
GGCTGGCCGAGTTCATTTATATAACAAACAAATATATAAAACAAGGGG
TAAAATGCCATCTTATATGATTAAACCGTCTCTTTTTTTTTTT
TTTTTTTTTTTTTTGTAATTAAAGAAGGGTATACCAGTGTCA
25 CTCTTATTCCAACCAGTCAAATAGGGACTTAGGTTGGAAACAGTT
CCGTGAGACCGTGAATTGGATGGTAGATAAATTAGTAAACTTAAACCTT
CAATTAAACCTACCTTTCTTATTAACCTCAATTCAACCTAAATTCTGAT
TCTTGGTTGAAAATAAGTTGCATCTTATTGTTGATTATCTGTCAT
AGGATCCTAGCATCTTTAATAATTATTGAAGGTGAAAGATCCA
30 ATTGAACTCTGTTGACGTTCCATCATTGCAACTGTTCTGAAAAAA
AAAATACCTAAATCAAATAACCATTCAAAATCCAAAATTATAAGAGA
GAATTGTAATGGACATGGAATCATAAAATCTTAAACACAGTTCA
AAGTTGCTAATTACATTCTGCTGTGCAGATTGAAATTCTATCAGAGAA
AGAGACATTACAAGAACGCCACTGGCAGTATTCAAATCTGATTCCC
35 CCTGTCTCATGCACCTTTCTACACCTCCGTGTGCTACATTGGATAAT
TATGAAGGAGTGGAGGTGGTGTGAGATAGAGAGTGTGAGAGTCCA
TAGAGAATTGGTAACAACTCACAATAACCAACAAACAGCCTATTACTTC
CCTACCTCCAGGAATTGTATCTAAGGAATATGGACAACACGAGTCATGT
TGGAAAGTGCAGCAACTGGAATAATTCTCACTCTCCAAAACA
40 AGAATCACCATTCCACAAACCTCACAACCATAGAAATGAGATGGTGT
GCTTTAGGTACTTGTGTTCGCCTCTCATGGCAGAACTCTTCAACCTA
AAGAAAGTCAAGATACTTGGGTGTGATGGTATTAAGAAGTGT
CAGAGATGATGAGGATGAAGAAATGACTACATTACATCTACCC
ACAAAAA

CCACCAACTTGTCCCTCATTTGATTCTCACTCTAAACCAACTGAAG
AATCTGAAGTGTATTGGTGGAGGTGGTCCAAGGATGAGGGGAGCAATGA
AATATCTTCAATAATACCACTGCAACGACTGCTGTTCTGATCAATTG
AGGTATGCTTGTACATATTCAATTATTATTAAATTCTTTTATT
5 GCAATATTCTATAAATAACATTATACCCACTATAACTAAGATAATAA
TTACCTAGAGGGATGGATGCTATGACACAGCTGCTACACTCAGAAACTC
TAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATTTGA
TGGGTAATATAGGCAATTAAAGTTTATTCTGTTAAAGCAGTATTAGC
AAGTACTGGCCAGTAGGAGAGGAATATCACCTTGAAAATCTGGT
10 CATTGTACCCAGAATTAGTTAAATGTAACATTAGATATTAGGGTTA
TCAGGTGACAGATATTGAGAATAGAACATATGTAATATTACCCAAAAC
TATTCTAAGGGTGCCTGTTAAATATGCTTCTGATTTCATTG
AATTGCAATTCTATATTAGGTGGTAAAGTGATTGCTCTCAATAAA
TCCCGAAATTAAATTAAAAAAAAACAAAAGTAAATTGGATATGGA
15 GAGCACTGGTATCATTAGTATATAAAAAACTAGATTGAAATTAGTT
TCTTATATAAAAGCTGTATAGTTAATTAGTTACATCATTTC
CATGTGGTGTGCAGTTGTCTGAAGCAGGTGGTCTTGGAGCTTATG
CCAATACGCTAGAGAGATAAAATAGGCAACTGCCATGCATTGCAAGTG
TGATTCCATGTTATGCAGCAGGACAATGCAAAAGCTCAAGTGCTGAGA
20 GTAATGGCTTGCAATGGGATGAAGGAGGTATTGAAACTCAATTAGGAC
GAGCAGCAACAAAAACAACGAGAAGAGTGGTGTGAGGAAGGAATTCAA
GAGTAAATAACAATGTTATTATGCTTCCAACTAAAGATATTAGTATT
GGAAATTGTGGGGTTGGAACATATATTCACATTCTGCACTTGAAAG
CCTGAGACAGCTCCAAGAGTTAACGATTAAGGGTTGCTACAGAATGAAAG
25 TGATTGTGAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACAACA
ACAACAACGAAGGGGCATCTTCTTCTTCTAAGAAGGTGGT
GGTCTTCTTGTCTAAAGTCCATTGATTGGTCAATCTACCAGAGCTGG
TAGGATTCTTCTGGGATGAATGAGTTCCGGTTGCCTTCATTAGATAAA
CTTATCATCGAGAAATGCCAAAAATGATGGTGTACAGCTGGGGTC
30 CACAGCTCCCCACTCAAGTATATACACACAAGATTAGGCAAACATACTC
TTGATCAAGAATCTGGCCTTAACITTCATCAGGTACATATATATTCTT
AATTGGCATCATCTAATTAAAGAAAGATATCATTCTGCCAAGTAAATTAA
CTTCAACACATTACACTAGTTCAAGTTATGTTCTAGGA
AGGCCAAATGGGAAAGCAAGATAGGGAAAAATAGAGTATTCAAGTGGAA
35 AGGGTATTCTAGGTATTCTGTCAAAATTGTTATTGCAAGGCTTTAG
TACCTGGAAGAGCATGATTATTCTCGATTGCTTCTTCTTATCATTT
TCTTAGCCTAGCATGATTTCATGAAATCTTCCCTGTTACTCCATTG
ATTGTTGTTCTCATGGTTCTAAGTGAGTTAGTGGCTCATCTGTTACTTC
TTTGATTGTTATTTCATAGCATGTTGTCATTGAATCAAGCTTTCCA
40 TTTCAACAAGGACAAAAGGTCAAAACTAACCTACTTATGAGATCAAA
ATAGCAACCCATCGGATAACTTTAGTTGGAGTTAAACTTACAATTACC
ATTGTGATTAATAATTATAATCTTGTATTAATTCAAAAAATTGGTAC
AGCACATATATGACATTCAAAGGTTTGTGACATATATGCCTCT

GGC GTTTCTTATTGGACATGCAGACTCATCCAAAGTTATACGGTG
 ACACCTTGGGCCCTGCTACTCAGAAGGGACAACCTGGTCTTTCATAAC
 TTTATCGAATTAGATGTGAAATCTAACATGATGTTAAAAGATTATTCC
 ATCCAGTGAGTTGCTGCAACTGCAAAGCTGGTAAAGATTAATGTAATGT
 5 GGTGAAAAGGGTAGAGGGAGGTATTGAAACTGCATTGGAAGCAGCAGGG
 AGAAATGGAATAGTGGATTGGTTGATGAATCGTCACAAACAACACTAC
 CACTACTCTTGTCAATCTTCAAACCTTGGAGAAATGAAGTTACGGGTC
 TCGATTGTCTGAGGTATATATGGAAGAGCAATCAGTGGACAGCATTGAG
 TTCCAAACCTAACAGAGTTGAAATTATGAATGTAATTCAATTAGAACAC
 10 TGTATTTACTAGTCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTAG
 AGATTGGTTGTGCAACCATATGGAGGTGCGATGTTCAGGATGCAGAT
 GTTCTGTAGAAGAACAAAGAGAAAGAACATCTGATGGCAAGATGAATAA
 GGAGATACTGTGTTACCTCATCTAAAGTCATTGAAATTACTACTTCTTC
 AAAGTCTTAAGGGGTTAGCTGGGAAGGAGGATTTTCACTCCCATTAA
 15 TTGGATACTTGGAAATCTACGAATGCCAGCAATAACCACCTCACCAA
 GGGAAATTCCGCTACTCCACAGCTAAAGAAATGGAACAAATTGGCT
 TCTTTATGTCAGGGAAAAAGACATCAACTCCTCTATTATAAAAGATC
 AAACAACAGGTAAACCAGATCTTGTGCTTAATAATTCTAAACTACA
 TTTGAAAAGCTTCATGCAAGTTTTTGTATATTGTCAAAACCGCAA
 20 CCTACATTTCAGCTTATATTATGTACTTTATGCAGGATTCAAACAA
 GACTCTGATTAATGTGAAGTGAATATTAAAGGTAAATTATATTTCATGT
 TCCTAGTTGCCTATTAAATTAAAGGCCTTTAGTCGTGATTTGGATGT
 ATTCTTCATGATGATGTCAATCTCTAAATACCCCATTCAATTGTTGGTTG
 AATGTTGACTCTATGTCAGGATGAATATTCAAGGGAAAGAATTGTTCATCA
 25 TATGAAGGACATTAAAGAACATGGTGCTAT

RG2C deduced polypeptide sequence (SEQ ID NO:92)

MAMETANEIIKQVVPVLMPINDYLRYLVSCRKYISDMDLKMELKEAKDNVEEH
 KNHNISNRLEVPAQVQSLEDVEKINAKVETVPKDVGCCFNLKIRYRAGRDAFNI
 30 IEEIDSVMRHSLITWTDHPIPLGRVDMSMASTSTLSTEHNDFQSREVRFSEALKALE
 ANHMIALCGMGGVGKTHMMQRLKKVAKEKRKFGYIIEAVIGEISDPIAIQQVVADY
 LCIELKESDKKTRAEKLQRQFKAKSDGGNTKFLIILDDVVWQSVDLEDIGLSPSPNQG
 VDFKVLLTSRDEHVCVSMGVEANSIINVGLIEAEAQRLFQQFVETSEPELHKIGEDI
 VRRCGGLPIAIKTMACTLRNKRKDAWKDALSRQLQHHIGNVATAVFRTSYENLPD
 35 KETKSVFLMCGLFPEDFNIPTTEELMRYGWGLKLFDRVYTIIEARNRLNTCIDRLVQT
 NLLIGSDNGVHVKMHDLVRAFVLGMYSEVEQASIVHGNMPGWPDENMDIVHSC
 KRISLTCKGMIEFPVDLKFPKLTIKLMHGDKSLKFPQEYEGMEKLRVISYHMKY
 PLLPLAPQCSTNIRVLHLTECSLKMFDCCSICGNLSNLEVLSFANSCIEWLPSTVRNLK
 KLRLLDLRLCYGLRIEQGVLKSLVKLEFYIGNAYGFIDDNCKEMAERSYNLSALEF
 40 AFFNNKAEVKNMSFENLERFKISVGCSFDGNIMMSSHSYENMLRLVTNKGDVLDISK
 LNGLFLKTEVLFLSVHGMNDLEDVEVKSTHPTQSSFCNLKVLIISKCVELRYLFKL
 NVANTLSRLEHLEVCKCKNMEELIHTGIGGCGETITFPKLKFLSQLPKLSQLCH

NVNIIGLPHLVDLKLKGIPGFTVIYPQNKLRRTSSLKEEVVIPKLETLQIDDMENLEI.
 WPCELSGGEVKLREIKVSSCDKLVNLPRNPMSSLHHLEELTVENCGSIESLFNID
 LDCVGAIGEEDNKSLRSINVENLGKLREVWRIKGADNSHLINGFQAVESIKIEKCK
 RFRNIFTPTANFYLVALLEIQIEGCGGNHESEEQIEILSEKETLQEATGSISNLVFPSC
 5 LMHSFHNLRLVTLDNYEGVEVVFEIESPTSRELVTTHNNQQPIILPYLQELYLR
 NMDNTSHVWKCSWNKFFTLPKQQSESPFHNLTTIEMRWCHGFRLFSPLMAELL
 SNLKKVKILGCDGIKEVVSNRDDEEEMTTFTSTHKTTLFPHLDSTLNQLKNLK
 CIGGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVWSLCQYAREIKIGNCHAL
 SSVIPCYAAQMOKLQVLRVMACNGMKEVFETQLGTSSNKNNEKSGCEEGIPRVN
 10 NNVIMLPNLKILSIGNCGGLEHIFTFSALESRLQLQELTIKGCYRMKVIVKKEEDEYG
 EQQTTTTTKGASSSSSSKKVVVFPCLSIVLNLPELVGFFLGMNEFRPLPSLDKLII
 EKCPKMMVFTAGGSTAPQLKYIHTRLGKHTLDQESGLNFHQTSFQSLYGDTLGPAT
 SEGTTWSFHNFIELDVKSNDVKKIIPSSELLQLQKLVKINVWMWCKRVEEVFETALE
 AAGRNGNSGIGFDESSQTTTTLVNLPNLGEMLRGLDCLRYIWKSQNQWTAFEPN
 15 LTRVEIYECNSLEHVFTSSMVGSLQLQLEIGLCNHMEVVHVQDADVSVEEDKEK
 ESDGKMNKEILVPLHLKSLKLLLQLSKGFLGKEDFSFPLLDTLEIYECPAITTFTK
 GNS.ATPQLKEMETNFGFFYAAGEKDINSSIICKQQDFKQDSD.

RG2D polynucleotide sequence (SEQ ID NO:93) and (SEQ ID NO:94)

20 ACGACCACTATAGGGCGAATTGGGCCGACGTCGATGCTCCGGCCGCC
 ATGGCCGCGGGATGTAACACGACGCCAGTCGAATCGAACCGTTCGTAC
 GAGAACATCGCTGCCTCTCCTCAACCATTAAATGTATATGAGCTAAATTG
 AAACATCTACTATCATGTTAAATTATAAACCTTTCTTAGATTAC
 TTGTCTGGATGTGTTAATAAAAACCAATTCCCACATCGTAGAGATCA
 25 TAGATGTAACTATTGTTAATCAATTTCGCTGCCAAGTTTAATAATTAT
 ACTTGGATATTAACAAAACCTTATCTAACGACCAAGGTAATATTAAAAAT
 AGGTTATTATTCTCATGCTAATTAAAAGATGGGTTGCAAAGTGAGACC
 ATGAAAACATTAACACGTTGATATTTCACACTTTATTCTTCATATTCA
 CCATATTTCGTTACTTCGTATTGATTAATCATCTTCATCACAGGGCTCC
 30 TTGGCAAAAGTCAGATCTATTAACAAACTTCCATGTGGTTGCAAATT
 ACAAGGATTCAACATAATTACCAAAACATAGCATTATCATAAGATCGAA
 TAATAATCAAATTCTCTATAATATTACACAAAGGTAACGTCTTAATTA
 ATTACGATACGAGACAGACTTTCACTCGTGACATCAACGGTCTATTCT
 AACCTTACTTAATTAAATGAATCTAGGATGTGCTCATATGCATGTAATAT
 35 TTGCTACCGTCATCTTCAAATGACCATATTGTTATGTATTATAATGAA
 TCAATGAAAAACCGGATTCTATTAAAATTCTTAAACTTCATCTTTA
 AGCCAGGGTGAATACAATTGTAGATCCACTGTTAATTCCATCGATTATG
 CGTGATCAATTGTTGGCTGCATACGATGCAGGTGCTACCACAAGAATATG
 GCCATGGAAACTGCTAATGAAATTATAAAACAAGTTGTTCCAGTTCTCAT
 40 GGTTCCATTAAACGATTACCTACGCTACGTCGTTCTGCAGAAAGTACA
 TCAGTGACATGGATTGAAAATGAAGGAATTAAAAGAAGCAAAAGACAAT
 GTTGAAGAGCACAAGAATCATACATTAGTAATCGTCTGAGGTTCCAGC

AGCTCAAGTCCAGAGCTGGTTGGAAGATGTAGAAAAGATCAATGCAAAAG
TGGAAACTGTTCTAAAGATGTCGGCTGTTCAATCTAAAGATTAGG
TACAGGGCCGGAAGGGATGCCTCAATATAATTGAGGAGATCGACAGTGT
CATGAGACGACACTCTGATCAGTGGACCGATCATCCCATTCTTGG
5 GAAGAGTTGATTCCGTGATGGCATCCACCTCTACGCTTCAACTGAACAC
AATGACTTCCAGTCAAGAGAGGTAAGGTTAGTGAAGCACTCAAAGCACT
TGAGGCCAACACATGATAGCATTATGTGGAATGGGAGAGTGGGAAGA
CCCACATGATGCAAAGGCTGAAGAAGGTTGCCAAAGAAAAGAGGAAGTT
GGTTATATCATCGAGGCAGTTAGGGAAATATCGGACCCCATTGCTAT
10 TCAGCAAGTTGAGCAGATTACCTATGCATAGAGCTGAAAGAAAGCGATA
AGAAAACAAGAGCTGAGAAGCTCGCAAGGGTCAAGGCCAAATCAGAT
GGAGGTAACACTAAGTCCCTCATAATATTGGATGATGTCAGTCCGT
TGATCTAGAAGATATTGGTTAACGCCCTCTCCAATCAAGGTGTCGACT
TCAAGGTCTGTTGACTCACGAGACGAACATGTTGCTCAGTGATGGGG
15 GTTGAAGCTAATTCAATTATTAACGTGGACTTCTAATTGAAGCAGAAC
ACAAAGATTGTTCCAGCAATTGTAGAAACTTCTGAGCCCGAGCTCCACA
AGATAGGAGAAGATATTGTTAGGAGGTGTTGCGGTCTACCCATTGCCATC
AAAACCATGGCGTGTACTCTAAGAAATAAAAGAAAGGATGCATGGAAGGA
TGCACTTCTCGTTACAACACCATGACATTGTAATGTTGCTACTGCAG
20 TTTTAGAACAGCTATGAGAATCTCCGGACAAGGAGACAAATCTGTT
TTTTGATGTTGTTGTTCCGAAGACTTCAATATTCTACCGAGGA
GTTGATGAGGTATGGATGGGCTTAAAGTTATTGATAGAGTTATACAA
TTATAGAAGCAAGAAACAGGCTAACACCTGCATTGAGCGACTGGTGCAG
GCAAATTACTAATTGGAAGTGATAATGGTGTACACGTCAAGATGCATGA
25 TCTGGTCCGTCTTGTGTTGGTATGTATTCTGAAGTCAGCAAGCTT
CAATTGTCAACCAGCTGTAATATGCCTGGTGGCCTGATGAAAATGATATG
ATCGTGCACCTTGCAAAAGAATTCTATTAAACATGCAAGGGTATGATTGA
GATTCCAGTAGACCTCAAGTTCTAAACTAACGATTGAAACTTATGC
ATGGAGATAAGTCTAAAGTTCTCAAGAATTATGAAGGAATGGAA
30 AAGCTCCAGGTTATATCATACTGATAAAATGAAGTACCCATTGCTCCTT
GGCACCTCAATGCTCCACCAACATTGGGTGTTCATCTCACTGAATGTT
CATTAAAGATGTTGATTGCTCTATCGGAAATCTATCGAATCTGGAA
GTGCTGAGCTTGCTAATTCTCGCATTGAATGGTACCTTCCACAGTCAG
AAATTAAAGAAGCTAAGGTTACTGATCTGAGATTGATGGTCTCC
35 GTATAGAACAGGGTCTGAAAAGTTGGTCAAACATTGAAGAATTAT
ATTGGAAATGCATATGGTTATAGATGATACTGCAAGGACATGGCAGA
GCGTTCTTACAACCTTCTGCATTAGAATTGCGTTCTTAATAACAAGG
CTGAAGTAAAAATATGTCATTGAGAATCTGAACGATTCAAGATCTCA
GTGGGGTGTCTTGTGATGGAAATATCAGTATGAGTAGCCACTCATACGA
40 AAACATGTTGCAATTGGTGACCAACAAAGGTGATGTATTAGACTCTAAC
TTAATGGGTTATTTGAAAACAGAGGTGCTTTTAAGTGTGCATGGC
ATGAATGATCTGAAAGATGTTGAGGTGAAGTCGACACATCCTACTCAGTC
CTCTCATTCTGCAATTAAAAGTCCGTATTATTCAAAGTGTGTAGAGT

TGAGATACTTTCAAACCTCATGTTGCAAACACTTGTCAAGCCTTGAG
CATCTAGAAGTTGTGGATGCGAAAATATGGAAGAACTCATACATACTGG
GATTGGGGTTGTGGAGAAGAGACAATTACTTCCCCAAGCTGAAGTCTT
TATCTTGAGTCACACTACCGAAGTTATCAGGTTGTGCCATAATGTCAAC
5 ATAATTGGGCCTACCACATCTCGTAGACTTGAAACTTAAGGGCATTCCAGG
TTTCACAGTCATTATCCGAGAACAGTTGCGAACATCTAGTTGTTGA
AGGAAGAGGTAGATATATGTTCTTATGTTAATACAATTAAATAATT
TTCAACCAAAATTTCATAATATATCTGTAATTGATTGTATGATGTGTTA
TTGTTATATGTGGCTATTAAAGGGATGATTATTGAGGTTGTGATTCC
10 TAAGTTGGAGACACTTCAAATTGATGGCATGGAGAACCTAGAAGAAATAT
GGCCTTGTGAGCTTAGTGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAA
GTGAGTAGCTGTGATAAGCTTGTGAATCTATTCCGACAATCCCATGTC
TCTGTTGCATCATCTGAAGAGCTTAAAGTCAAAATTGTCGTTCCATTG
AGTCGTTATTCAACATCGACTTGGATTGTGTCAGTGCAATTGGAGAAGAA
15 GACAACAAAGAGCATCTTAAGAAGAACAAAGTGAAGAATTAGGGAAAGCT
AAGAGAGGTGTGGAGGATAAAAGGTGCAGATAACTCTCGTCCCCTCATCC
ATGGCTTCCAGCTGTGAAAGCATAAGTATCTGGGGATGTAAGCGGTTT
AGAAATATATTACACACCTATCACCGCCAATTGATCTGGTGGCACTTTT
GGAGATTCACATAGGAAATTACAGAGAAAATCATGAATCGGAAGAGCAGG
20 TAACGCTTCAATTCACTTCTTACTTAATTAAAGGACTAAGCTTGTGTT
TTTGATAATAAAAGAGGTGGGATGACTAAACTGGGATCACAATTGTA
ACAAAATGTTACAAACCATGAACGTACAAACCATTCTGAATTAGGTT
TCAATACAAGTCATTACAAATATGGCTTAAGTTTTTATATTATGTA
TCAACATTATTTCTTACAGAGGTATTATTATAATAGTAAGTTAAAGC
25 AATTAAATTAGCACTAATTTCATCATCTAACTTAGCTAATAATCG
TTATAAAATGTCAATAGCTAAAATAATTGACATTCACTGAGAGCA
ATTTTTCTAAACATGATTGCAAATGATTAAAACCTTAAATTAAACTAAA
AAGATTTTATATATGTTACAAAATTACAAATTGAAATTGGATATGT
TAATTAAACAGTTATAATTATTGTATTACAAAGCGATATATAATAAAAATA
30 TTATTTCTGTAGTCATGTATAATTGTATATGAAATGATTTTAAGA
TGGTAGAAGTGGAAACTAGTCATCTCACTTAACCTCATGTCACACCA
TTTATATCCGTTCTCTCTCTCTCTGCTCCATCTTTCAAC
TCATAACACATAAAAATAACATATTCCAACACATTAAAGTCACTACCA
CATCATTATTTAATTAAATTAGAAAATATAAAATTAAATAAAA
35 CATAACATTTTTATTAAAGGCACTAATACAATAAAAAGATAACACGG
TAAATAAAAAAACGATAATTAGAAAAAAACATAATAAAAAAGACAACA
TTA AAAATAWAAGCGACAACCTAAATTAACTAATGATCAAGAAAATTCT
AAAACCCACCATATTCTGCAATTGTCATTATGTTCAAACACCA
TTCGCAGAACCTCCCTATCAAGTGTATCTCCAAGTCAACAAAGTTAGCTTCAAC
40 ATGTCTCTCATGTATCTCCAAGTCAACAAAGTTAGCTTCAATTCTC
ATTTCATGTAAGACGCAAATTTCATCCGATATTGTTCTATCTT
CCACCTCTACTTATTACACAGTGTGGATGAAGGGAGAGCACAGCGATTCTC
GTACGAACGGTTACGATTGACTGGCCGTCGTTACAATCCCGCGGCCA

TGGCGGCCGGGAGCATGCGACGTCGGGCCATTGCCCTATAGGGTCGT
AATACA (SEQ ID NO:93)

Sequence gap

5 TGAGCCTCCGATGCTTAGTCCACTTGGCACAGTTCAAGTCCAATCAACTT
ATAACCCATTTCCTCAAGTTGTCTCAAGTTAAGGCCAATTGCCCTC
TCCAAATCATCCATAACTTCATGGAATGCCCTTCATCTTAATCCGAA
TGCACAATTATTCTCCATCTCATTTAAGCAAGAGGCCACCTCTTC
TGCTTCATCCATCAATAGTCTGTTGAATAGTGTCTAAGGCTGCAACTAT
ATTAGACAAGTATTGACCCGGTTGTGCATGGCCTTTGGGTTGCCTTC
10 ACCATAGCAACTTGATAGGATGATTATTAAAGAGAGATAATATTATA
ATATATTATGAGAATAATATAATGAATAATATATTGTATTGATTAAT
ATAAGTCATAGAATTAAATTAGAATTAAATTGGTGACTTAAAGAGATTAAT
TAAATAAAGGGTATAAACGTCAATTGTTGATAGTTAAGCTTAGACT
GTAATCCATTGGATATGGTATGGACGAATCCTAAGGGATTAGGATAG
15 CTAATCGTCATATGAGTTATCTAAGAAGGATTGGATAGCCTTAAGA
GAAGATTATCTGATAGGGACTTATCTGTAATCCTTAAGGAGTCTACAAGT
ATAAAATAGACCCTATGGCTGATGGAATTCGACACATCTCCTAAAGTAAGA
GAGCCTTGGCGAATTCCCTCCCTCACCTCTCCTAAATCATTCTT
GCTATTGGTGTGTTGAAGCCATTAGAGGAGTGACATTGTGACTCTAGAA
20 TCTCCAAGACCTCAAGATCAACAAGGAATTCAAAGGTATGATTCTAGATC
TGTTCATGTTATTGCTTAATTAGTCATTAGAAGACTTGGATTC
AAAGCATGTTATTAGAAAGCCTAGATCYGAGCAATAGGGTTGATGC
GCACATAGGAAAGTCTTATGGCTAAAACCCATCATAGCCACTTCATGT
ATCATCTCTACTAGTTATTAGCCATAATCCTGTTGCTCTCCAAGTT
25 AATTACCTCCCTAGTCCTGTTCTGCTAGTTCTTAAATTGCTATT
AAGATCACAGAACTAGAGAGTACCCAAAATGGTTATAAAATAACAAAAAG
GAAATATGCATGAAGATTAACAAATTATAATGTAATATGCTAAAATA
AACTATAAAAAAAAGTAAATAAAATGAAACTATCACACTCCGACCACCC
TTATAGGCTGTACTGCACCCACCTCATTCTGTACCAATATGGGAT
30 GGAAACATTATTCATTAAGCCAAAAACTAACATTAAGGGTGAGTGAC
AAAGGTAAACTAAAGACAACAATAATCCATTCTTCTGTACATACACA
ACACACACATAGGGCGGACGTAGGATTGAGTATGTGTTGCTGGGTGAC
ACATTCTTCTTACGTAGTACACAATAGTAGAGAAAACGAGAAATTC
CAATTCTTACATTGTGTTGAAAAATATAACAGGGTTGCTGGTGTAC
35 TCTGGGCACCAAGTGGAACCGCCCTGCACACACACACATAGAGGGA
GAGAGAGAGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATT
TGGGATGTGATACTCTTGGAAAATGGAGCAATATCTTAATATTGT
ATTTTTTAATGTAATTATATAATTAACTTACATTAGTTATAACTTTA
GTTCCTTATTAACTGTATATTAAATCATTCAAGTTATAAGTTT
40 ATTATTTGGTATACCAGAAAAAAAGTCTTATGTGTTGGATTAAAC
ATAAAAATCTAACAAATTAAATCAAAAAGACCAACATGTGGACAATTAT
GTATATAATTCTCAATGGCTTAGTGTAAACGATATAAATTCAAAA
CAATTCTCACATTAAAAACACTTCAGTCATAATTGTTATAAATTA

TCATTGTATCACAAAATCAGTTCATACATCACATCCCAAGATCAATAAA
GTGTAAATACTCCTCATGTGTACTAATCAAGCCGACGCCCTCCCGGA
TTCTCACTGGTACCTGAAACACGTAACATAACAACGTAAAGCATAATGC
TTAGTGAGTCCCCAAAATACCACATACCACATATGCCTTCAGGCC
5 ATAACCTCTGTAGGATCTCCGACCCAAAGTGTCTCAGGGGACTCCGTCCC
GAATCCCGGTAGACCTCCGGTCTACCGTATTGACCTCCGGTCCGTA
TCATACATAACATAACACATACATACATAACAACATATAGCAC
ATACATCTCATAACATAAAAGACCTCCGGTACACATAAGGTACCCCTCC
AGGTACAGTATAGTGAGAANACTCACCTGTATGATGTCTAACCTCAC
10 GTGCTCGATATCCCTGAATCTCGAAACAAATGACCTAGCCCCGCCTACTCA
CATAAAGTAATTATTCAAATCATTAACGGCTCTCAAGGCTAGACTACAT
CCCTTCTATAAATCCACAGAAGGGTAAAGACCACTTACCCCTCTTG
ACCCAAAAGTCAAATGTTGATCAAACCCCCAAAAGTCAACGAAAGACAA
TGGTCAACTTGTACCCCTACTCGTGGAGTGCACAAAGGTGACTCGGCAAGT
15 ACATGCGGGTCCTCTGAATCTTCAGTCTCTTGGCTCGTAGTCTT
TCTTCCACCCGACGAGTTACACCTGTCTGAATCGCGGGCAACCCCCGAC
TCGACTTGTGAGTCCGCTCATGGACTCAACGAGTTCAATTCCATGCTCAC
ACTCAAATGACCTCTGAGGTCAGATCTGTTCTCTAAATCCATAGATCTG
ACCTTCCAAGCTCAATAAACACGTAAAGGTTGAACTTGATACTCATGC
20 AACGTCCAATGATTCTACTTGATGATTAGCCCCAAATACAACATCCTA
AGTCCATACGACCTTATTTCTCAAATAACAAACACATATATTAATTAC
CAATGACAGTAATAGATATCATATAAAGTATTGTAACACTTGTAAAGAA
CCTTGCTACTATAGGTAAGAACATTTCAAAGTACATGCCCTAATT
GAAAAAAAGTTATAAAAAAATAATGACTAGGGCGTGTGTTTTACTAG
25 TTTGTATCAAATTATATCAAATTAAGGTGAAAAGAATGACGACCACA
TTAACCGAAAATGTAATTATTTTTATTGTAATTGTTAATTGTTAATTGTT
GTGATCTATGTATTAAAAGTAAATATCAAACAAAGAACATAATCCAAACC
CTAAATTGCAAGTCTCGCCCAATTCTCTATCACTAGTCCTCACTTACGA
TGGCGTTACGTCGCTCTCACTTCCACAACCCATTGTTGCTACTAATT
30 ACACTAACGAAAAGTTGAATATCCATATATTATTGATGTGAAATTGA
ACGAATCTCGTCAAATTATTATTGATGGATTGAGTGGAAAGTT
AGGCAGAACGGGAATGATGGTCTGCAAGTGGTTATAAACATGGGTGAAGA
TAAAATGGAGTTGCGCCGGTGTATTATAGATCTCTTGGGTTGATT
TGAGTTATTACTGTATACGTAGCCTCTTACAACGACCATTCTCCAAGT
35 ACCATTGATCTTTAGAATCCAGTTGTCTGAAACACCCCTGATTGGAT
CAAATATCACCAACAACCTTAAGAACTGGACTAATTAAATTGTTCTTG
ATCTTGATAACAAGAGGAAACACGTCACCATATCTTATTAAATTG
CTTTGGTGTATTCTTCTTCCATTCTTGATCTGTTCCAGAT
GGTATTGGTGTGGATAATTACACCTGGAGATTGTGAACGATGGGAAGG
40 GGTATGTGATTACAGAGGATGTGGCTGTGGTTGAGGATGGTTATGGC
TGGCGAGTCTAATTATTTATATAAACAAATAATATAAAACAAG
GGTAAAATATGTATTAAAGCGCCTTAAATGGTGACAATTTCACAG
TTTACTCTTTGTTTTAATTGTGATGCCACGATCGAACTCATTCTA

CCCCCCCCCTTTTTTAAAATAAAAAATTAAAGAAGGGTACCAACCAT
ATACCCGTGTCAGCTCTTATTCCCAAGCAGTCAAATAGGGACTTAGGTT
GTATGGAAACAGTCCGTACTGGATGGCAGATAAATTAGTAAACTTA
ACCCTCAATTAAACCTACCTTTCTTATTAACTCAATTCAAGCTAAAT
5 TCTGATTCTGTTGAAAATAAGTTGCATCTTATTGGCATATTATCT
TGGCATAGGATCCTAGCATCTTAATAGTTATTGAAGCTGAAAG
ATCCAACTAGTTGATCTGGCATTTCCATCATTGCAACTGTTTC
TTGAAAAAAAATACCTAAAATCAAATAACCATTTCAAATCCAAAATTA
TAAGAGAGAATTGTAATGGACGTGGAATCATAATCATTAACACAGTTC
10 AGTACACAAGTTGTAATTACATTCTTGCTGTCAGATTGAAATTCTAT
CAGAGAAAGAGACATTACAAGAAGTCACTGATACTAATATTCTAATGAT
GTTGTATTATTCCCACCTGTCATGCACTCTTCATAACCTCCATAA
ACTTAAATTGAAAATTATGAAGGAGTGGAGGTGGTGGAGATAGAGA
GTGAGAGTCCAACATGTAGAGAATTGTAACAACTCACAATAACCAACAA
15 CAGCCTATTATACTCCAACCTCCAGGAATTGTATCTAAGGAATATGGA
CAACACGAGTCATGTGGAAGTGCAGCACTGGAATAATTCTCACTC
TTCCAAAACAACAATCAGAACATCACCATTCCACAACCTCACAAACCATAA
ATGAGATGGTGTATGGCTTAGGTACTGTTTCGCCTCTATGGCAGA
ACTTCTTCCAACCTAAAGAAAGTCAGATACTGGGTGTATGGTATTG
20 AAGAAGTTGTTCAAACAGAGATGATGAGGATGAAGAAATGACTACATT
ACATCTACCCACACAACCACCAACTGTTCCCTCATCTGATTCTCAC
TCTAAAATACATGCACTGTGAAGTGTATTGGTGGAGGTGGTCCAAGG
ATGAGGGGAGCAATGAAATATCTTCATAATACCACTACAACCTACCGAT
CAATTAAAGGTATGTTGTACATATTAAATTATATTAAATTCTTGT
25 TAATTCCCTTCTTGCAATATTCTATGCGAACTCAAGAATGGGATTG
GAGGCATATAAAGTTACATTCAAGTACCTATAACATTCTTTATTGTT
ATTATCATTTCATATCAAGTACCTATAACATTCTTTATTGTT
AATTAGAAGAGGTCACATGTCTAATTAGGTTCCATTCTATGTGAAC
CTCTATTCTCTGTAACTCAAGCATCTAGATTATTATCCATTTCATA
30 ATTGTGTTATTTCACAGTTTTTTATTAAATTAAATAATTAA
TTTAATTATTATTATTATTGTTGGTAATTGCAACCTGTATAT
TCAAGTCTTAATGTAACATAATAACATTTATACCCACTATAACTAAGA
TAATAATTACCTAAAGGGATGGATGCCATGACACTGCTACACTCAGNAA
CTCTAGTAAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTT
35 TGATGGGTAATATAGGCAATTAAAGTTATTCTGTAAAGCAGTATT
AGCTAGTAGTGGCCAGTAGGAGAGGAGAATATCACCTTGTCAAATCT
GGTCATTGTACCCAGAATTAGTTAAATGTAACATTAGATATTAGGG
TCATCAGGTGACAGATATTGAGAATAGAACAAATATGTAATATTACCAA
AACTATTTCATAAGGTTGCTCTGTAAATATGTGCTTCTGATTCA
40 TTGAATTGCAATTGCTATATTAGGTGGTAAACTGATTGCTCTTCAT
AAATCCTGAAATTAAATTAAAAAAAAACAAAGTACATTGATT
GGAGAGCACTGGTATCATTAGTATAGAAAAAAACTAGATTGATTAY
CTTCTTATATAAAAGTTGTATATAGTTAATTAGTTACATCATT

TTCTATGTGTTGCAGTTGCTGAAGCAGGTGGTGGAGCTT
 ATGCCAATACTCTAGAGAGATAGAGATATAGGTGTGACTGTCAA
 GTGTAATTCCATGTTACGCAGCAGGACAATGCAAAAGCTGCAAGTGCTG
 ACAGTCAGTCTTGTAAATGGCTGAAGGAGGTATTGAAACTCAATTAGG
 5 GACGAGCAGCAACAAAAACAAACGAGAAGAGTGGTTGTGAGGAAGGAATT
 CAAGAGTAAATAACAATGTTATTATGCTTCCAATCTAAAGATATTGGAA
 ATCTACGGTTGTGGGGTTGGAACATATATTCACATTCTGCACTTGA
 AAGCCTGAGACAGCTCCAAGAGTTAACGATTAAGGGTTACTACTCTGTC
 AATCTTCAAACCTCAAAGAAATGAGGTTGGAGTGGCTAAGTAATCTGAG
 10 GTATATATGGAAGAGCAATCAGTGGACAGCATTGAGTTCCAACCTAA
 CAAGAGTTGAAATTGTGAATGTAATTCACTAGAACATGTATTACTAGT
 TCCATGGTTGGTAGTCTATTGCAACTCCAAGAGCTACATATATTAACTG
 CAGTCTGATGGAGGAGGTAATTGTTAAGGATGCAGATGTTCTGTAGAAG
 AAGACAAAGAGAAAGAACATCTGATGGCAAGACGAATAAGGAGATACTTGTG
 15 TTACCTCATCTAAAGTCCTGAAATTACAACCTCTCGAAGTCTTAAGGG
 GTTAGCTGGGAAGGAGGATTTCAATTCCATTATTGGATACTTAG
 AAATCAAAGATGCCAACAAATAACCACCTCACCAAGGAAATTCCGCT
 ACTCCACAACAAAAGAAATACAAACAAATTGGCTCTTTATGCTGC
 AGGGGAAAAAGACATCAACTCTTATAAAGATCAAACAAACAGGTAAATC
 20 AGATCTTGTGCTTAATAATTCTAAACTACATTGAAAAGCTTCATG
 CAAGTTTTGTATATTGCAAAAACCGCAACCTACATTCACTGTTAT
 ATTATGTAATTGAGGATTCAAACAAGACTCAGATTAATGTGAAG
 TGAATATTAAAGGTAATTATATTTCATGTTCTAGTTGCCTATTAAATT
 AATGGCCTTTAGTCATGATTTGGATGTATTCTCATGATGATGTGA
 25 ATCTTCTAATACCCCATTCAATTGTTGGTGAATGTTGACTCTATGTCAG
 GATGAATATTCAAGGGAAGAATTGTCATCAWATGAAGGACATTAAAGAA
 CATGGATGCTATGAAGATGTTGGAAAACATATGTATCAAGTGGCAARCT
 GCTTAATGATCTAAGTTGTTGGTGANAGATGTTGATTAAATATTCAA
 ATTCAATTGGTATATGGCTTATCAATAGTGTAAATGGATAATGAGTGA
 30 CTTAACCTAAATTATGTTGGTAAATGTTGGACAAGTATGGAAAATTA
 GGAATGACTTGTGAAAAAAAATAAAAAAAA (SEQ ID NO:94)

RG2D deduced polypeptide sequence (SEQ ID NO:95)

MAMETANEIIKQVVPVLMVPINDYLRYVVSCRKYISDMDLKMKEAKDNVEE
 35 HKNHNISNRLEVPAAQVQSWLEDVEKINAKVETVKPDVGCCFNLKIRYRAGRDAF
 NIIEEIDSVMRRHSLITWTDHIPIPLGRVDSVMASTSTLSTEHNDQSREVRSEALKA
 LEANHMIALCGMGRVGKTHMMQRLKKVAKEKRKFGYIIEAVIGEISDPIAIQQVVA
 DYLCIELKESDKKTRAEKLRQGFKAQSDGGNTKFLIILDDVWQSVDLEDIGLSPSPN
 QGVDFKVLLTSRDEHVCVSMGVEANSIINVGLLIEAEAQRLFQQFVETSEPELHKIG
 40 EDIVRRCCGLPIAIKTMACTLRNKRKDAWKDALSRQLQHHDIGNVATAVFRTSYENL
 PDKETKSFLMCGLFPEDFNIPTTEELMRYGWGLKLFDRTVYTIIEARNRLNTIERLV
 QANLLIGSDNGVHVKMHDLVRAFVLGMYSEVEQASIVNHGNMPGPDENDMIVH

SCKRISLTCKGMIEIPVDLKFPKLTLKLMHGDKSLKFPQEYEGMEKLQVISYDKM
 KYPLLPLAPQCSTNIRVLHLTECSLKMFDCCSIGNLSNLEVLSFANSRIEWPSTVRN
 LKKLRLLDLRFCDFGLRIEQGVLKSLVKLEFYIGNAYGFIDDNCNDMAERSYNLSA
 LEFAFFNNKAEVKNMSFENLERFKISVGCSFDGNISMSSHYENMLQLVTNKGDVL
 5 DSKLNGLFLKTEVLFSLVHGMNDLEDVEVKSTHPTQSSFCNLKVRIISKCVELRYL
 FKLHVANTLSSLEHLEVCGCENMEELIHTGIGGCGEETITFPKLKSLSQLPKLSGL
 CHNVNIIGLPHLVDLKLKGIPGFTVIYPQNKLRTSSLKEEVVIPKLETLQIDGMENL
 EEIWPCESGGEKVKLREIKVSSCDKLVNLFPHPNPMSSLHHLEELKVKNCRSIESLF
 10 NIDLDCVSAIGEEDNKSILRRIKVKNLGKLREVWRIGADNSRPLIHGPAVESIWI
 GCKRFRNIFTPITANFDLVALLEIHIGNYRENHESEEQIEILSEKETLQEVTDTNISND
 VVLFPSCLMHSFHNLHKLKLENYEJVVEVVFIESESPTCRELVTTNNQQPIILPN
 LQELYLRNMDNTSHVWKCSWNKFFTLPKQQSESPFHNLTTIEMRWCHGFRYLF
 PLMAELLSNLKKVKILGCDGIEEVVSNRDDEDEEMTTFTSTHTTNLFPHLDSTLK
 YMHCCLKCIGGGAKDEGSNEISFNNTTTTDQFKLSEAGGVCWSLCQYSREIEIYRC
 15 DALSSVIPCYAAGQMQLQVLTVSSCNGLKEVFETQLGTSSNKNNEKGCEEGIPR
 VNNDVIMLPNLKILEIYGCGLLEHIFTSALESRLQLQELTIKYTLVNLPNLKEM
 RLEWLSNLRYIWKSQNQWTAFEFPNLTRVEICECNSLEHVFTSSMVGSLQLQELHIF
 NCSTMEEVIVKDADVSEEDKEKESDGKTNKEILVPLHLKSLKLQLRSLKGFSLGK
 EDFSFPLLDTLEIKRCPTITTFTKGNSATPQLKEIQTNGFFYAAKEKDINSLIKIKQQ
 20 DFKQDSD.CEVNIK

RG2E polynucleotide sequence (SEQ ID NO:96)

TGGGAAGACACAATGATGCAAAGGTTGAAGAAGGTTGCTAAAGAAAATAGAAT
 GTTCATTATATGGTTAGGCAGTTAGGGAAAAGACAGACCCACTTGCTAT
 25 TCAACAAGCTGTAGCGGATTACCTTGTATAGAGTAAAGAAAGCACTAAACC
 AGCAAGAGCTGATAAGCTTCGTGAATGGTTAACGCCAACTCTGGAGAAGGTA
 AGAATAAGTTCTTGTAAATTGATGATGTTGGCAGTCCGTTGATCTGGAAG
 ACATTGGTTAACGTCAATTCCAATCAAGGTGTCGACTCAAGGTCTTGTGA
 CTTCACGAGACGAACATGTTGCACAGTAATGGGGTTGAAGCTAATTCAATT
 30 TTAATGTGGACTCTAGTAGAACAGCAGAACAGCACAAAGTTGTTCCAGCAATTG
 TAGAAACTTTGAGCCCAGCTCCATAAGATAGGAGAACATCGTAAGGAAG
 TGTTGTGGTTACCTATTGCCATTAAAACCATGGCATGTACTCTAAGAAATAAA
 AGAAAGGATGCATGGAAGGATGCACTTTGCAATTAGAGTACCATGACATTAGC
 AGTGTGCGCCAAAGTCTTGAACAGAGCTACCATATCTCCACAACAAGGAG
 35 ACTAAATCTGTGTTTGTGATGTGGTTTTCTGAAGACTTCATATTCAA
 TCGAGGAGTTGATGAGGTATGGATGGGCTTAAAGATATTGATAGAGTTATA
 CTATTAGACAAGCAAGAACATCAGGCTAACACACCTGCATTGAGCGACTGGTGCAG
 ACAAAATTGTTAATAGAAAGTGTGATGGTGTGCACGTCAGATGCATGATCTG
 GTCCGTGCTTCGTTGGTTATGTTCTGAAGTGAACATGCTTCATTATCA
 40 ACCATGGTAATATGCTTGGATGGCCTGAAAATTATGACCAACTCTTGCAGAA
 CAATTCAATTACATGCAAGAGTATGTCTGAATTCCGGGAGATCTCAAGTTTC
 CAAACCTAACGATTGAAACTCATGCATGGAGATAAGTTGCTAAGATATCCTC

5 AAGACTTTATGAAGGAATGGAAAAGCTCTGGTTATATCATATGATGAAATGA
 AGTATCCATTGCTCCCTCGTTACCTCAATGCTCCATCAACCTCGAGTGCTCA
 CCTCCATCGATGCTCATTAATGATGTTGATTGCTCTGTATTGAAATA GTTG
 AATCTGGAAGTGCTTAGCTTGTAAATCTGGCATTGAATGGTACCTTCACA
 ATAGGAAATTAAAGAAGCTAAGGTTACTTGATCTGAGAGAGATTGTATGGTCTT
 CGTATAGAAAAAGGTGTCTGAAAAATTGGTAAAATTGGAGGAATTATATT
 GGTAGAGCAGATATTTATAGAT

RG2E deduced polypeptide sequence (SEQ ID NO:97)

10 WEDTMMQRLKKVAKENRMFNYMVEAVIGEKTDPLAIQQAVADYLCIELKESTKP
 ARADKLREWFKANSSEGKKNKFLVIFDDVWQSVDLEDIGLSHFPNQGVDFKVLLTS
 RDEHVCTVMGVEANSILNVGLLVEAEAQSLFQQFVETFEPELHKIGEDIVRKCCGL
 PIAIKTMACTLRNKRKDACKDALLHLEYHDISSVAPKVFETSYHNLHNKETKSVFL
 MCGFFPEDFNIPIEELMRYGWGLKIFDRVYTIRQARIRLNTCIERLVQTNLLIESDDG
 15 VHVKMHDLVRAFVLVMFSEVEHASIINHGNMLGPENYMTNSCKTISLTCKSMSE
 FPGDLKFPNLTILKLMHGDKLLRYPQDFYEGMEKLWVISYDEMKYPLLPSLPQCSI
 NLRVLHLHRCSSLMMFDSCSCIGNMLNLEVLSFKSGIEWLPSTIGNLKKRLLLDLRD
 CYGLRIEKGVKNLVKIGGIYIGRADIL.

20 RG2F polynucleotide sequence (SEQ ID NO:98)

CTGTGGAAGACACAATGATGCAAAGGCTGAAAAGGTTGTGCATGAAAAGAAA
 ATGTTAACCTTATTGTTGAAGCAGTTAGGGAAAAGACAGACCCCCGTTGCC
 ATTCAAGGATGCTATAGCAGATTACCTAGGTGTAGAGCTCAATGAAAATCTAAG
 CAAGCAAGAGCTGATAAGCTCCGTCAAGGATTCAAGGACAAATCAGATGGAGG
 25 CAAAATAAGTTCTTGTAAATTGACGATGTTGGCAGTCTGTTGATCTGGA
 AGATATTGGTTAACGTCTTCCAATCAAGCGTCGACTTCAAGGTCTTGT
 GACATCACGAGACAGACATGTTGCACAGTGTAGGGGGTTGAAGCCAATTAA
 TTCTAAACGTGGGACTCTAATTGAAGCTGAAGCACAAAGTTGTTCCACCAAT
 TTGTTGTCACCTCTGAGCCCAGCTCCATAAGATAGGGAGAAGATATTGTAAAGA
 30 AGTGTTCGGTCTGCCAATTGCCATCAAACCATGGCATGTACTCTACGACATA
 AAAGAAAGGATGCATGGAAGGATGCACTITCACGTTAGAGCACCATGACATT
 CAAAGTGTGCTAAAGTATTGAAACGAGCTACAACAATCTCAAAGACAA
 GGAGACTAAATCCGTATTTGATGTGGTTGTTCTGAAGACTTGGATAT
 ACCTATCGAGGAGTTGATGAGGTATGGATGGGGCTTAAGATTATTGATAGAGT
 35 TAACTATTACACAAGCAAGAAACAGGCTCAACACCTGCATTGAGCGACTGG
 TGCAACACAAATTGTTAACATGAAAGTGTGATGGTGTGCATGTCAAGATGCATG
 ATCTGGTTCTGCTTTGTTGGGAATGTTCTGAAGTGGAGCATGCTCAAT
 TGTCAACCATGGTAATATGCCGAGTGGACTGAAAATGATATGACTGACTCTG
 40 CAAACAAATTCTTACATGCAAGAGTATGTTGGAGTTCTGGAGACCTCAA
 GTTTCCAACCTAAAGATTGAAACTTATGCATGGAGGTAAGTCACTAAGGTA
 TCCTCAAGACTTTATCAAGGAATGGAAAAGCTGGAGGTTATATCATACGATGA
 AATGAAGTATCCATTGCTCCCTCGTTGCCTCAATGTTCCACCACCTTCGAGTG

CTTCATCTCCATGAATGTTCATTAAGGATGTTGATTGCTCTCAATCGGTAACTC.
 TTTCAACATGGAAGTGCTCAGCTTGCATAATTCTAGCATTGAATTGTTACCTTC
 CGTAATTGAAATTGAAGAAGTTGCGGCTGCTAGATTGACAAACTGTTATGG
 TGTTCGTATAGAAAAGGATGTCCTGAAAAATTGGTAAACTGAAGAGCTTA
 5 TATTAGGAATGGTCTACCAGTTACAGAGGAT

RG2F deduced polypeptide sequence (SEQ ID NO:99)

VEDTMMQRLKKVVHEKKMFNFIVEAVIGEKTDPAIQDAIAIDYLGVELNEKSKQA
 RADKLRQGFKDSDGGKNKFFVILDDWQSVDLEDIGLSPFPNQGVDFKVLLTSRD
 10 RHVCTVMGVEAKLILNVGLLIEAEAQSLFHQFVVTSEPELHKIGEDIVKKCFGLPIAI
 KTMACTLRHKRKDAWKDALSLEHHDIQSVPVKVFETSYNNLKDKEKSVFLMCG
 LFPEDLDIPIEELMRYGWGLRLFDRVNTITQARNRLNTCIELVHTNLLIESVDGVH
 VKMHDLVRAVLGMFSEVEHASIVNHGNMPWTENDMTSCKQISLTCKSMLEFP
 15 GDLKFPNLKILKLMHGGKSLRYPQDFYQGMEMKLEVISYDEMKYPLLPSLPQCSTILR
 VLHLHECSRMRMFDCSSIGNLFNMEVLSFANSSIELLPSVIGNLKKLRLLDLTNCYGV
 RIEKDVLKNLVKLELYIRNGLPVYRG

RG2G polynucleotide sequence (SEQ ID NO:100)

GAAGACACGATGATGAAGAACTGAAGGAGGTCGTGGGACAAAGAAATCATTCA
 20 AATATTATTATTCAAGTGGTCATAGGAGAGAACAAACCTATTGCAATTCTAG
 CAAGCTGTAGCAGATTACCTCTATAGAGCTGAAAGAAAACACTAAAGAAC
 AAGAGCTGATAAGCTCGTAAACGGTTGAAGCCGATGGAGGAAAGAATAAGT
 TCCTTGTAAACTTGACGATGTATGGCAGTTGTCGATCTGAAGATATTGGTT
 25 AAGTCCTCTGCCAATAAAGGTGTCAACTTCAAGGTCTTGTGACGTCAAGAGA
 TTCACATGTTGCACTCTGATGGGAGCTGAAGCAAATTCAATTCTTAATATAAA
 AGTTTAAAAGATGTAGAAGGACAAAGTTGTCGCCAGTTGCTAAAGATTC
 GGGTGATGATGACCTGGATCCTGCTTCAATGGATAGCAGATAGTATTGCAAG
 TAGATGTCAAGGTTGCCATTGCCATCAAACCATGCCCTAAGTCTAAAGG
 30 TAGAAGCAAGTCTGCATGGGACGTTGCACTTCTGCTGGAGAATCATAAGAT
 TGGTAGTGAAGAAGITGTGCGTGAAGTTTAAAATTAGCTACGACAATCTCCA
 AGATGAGGTTACTAAATCTATTTTTACTTTGTGCTTATTCTGAAGATT
 GATATTCTACTGAGGAGTTGGTGGAGGTATGGTGGGGCTGAAATTATTATA
 GAAGCAAAACTATAAGAGAACAGCAAGAACACAGGCTCAACACCTGACTGAGCG
 35 GCTTAGGGAGACAAATTGTTATTGGAAAGTGTGACATTGGATGTGTCAGAT
 GCACGATGTGGTGCCTGATTTGTTGCATATATTCTCAGAAGTCCAACACGC
 TTCAATTGTCAACCATTGTAACGTGTCAGAGTGGCTAGAGGAAATCATAGCAT
 CTACTCTGTAAAAGAATTTCATTAAACATGCAAGGGTATGTCTCAGTTCCCAA
 AGACCTCAAATTCCAAACCTTCAATTGAAACTTATGCATGGAGATAAGTC
 ACTGAGCTTCTGAAAACCTTATGGAAAGATGGAAAAGGTTCAGGTAATATC
 40 ATATGATAAAATTGATGTATCCATTGCTCCCTCATCACTGAATGCTCCACCAA
 CGTTCGAGTGCTTCATCTCATTACTGTTCAAGGATGTTGATTGCTCTCA
 ATTGGTAATCTTCTCAACATGGAAGTGCAGCTTGTCAATTCTAACATTGAA

TGGTTACCATCTACAATTGAAAGTTGAAGAAGCTAAGGCTACTAGATTGACA
 AATTGTAAAGGTCTCGTATAGATAATGGTGTCTAAAAAATTGGTCAAACCT
 GAAGAGCTTATATGGGTGTTAACGTCCGTATGGACAGGCCGTTAGCTGACA
 GATGAAAA

5

RG2G deduced polypeptide sequence (SEQ ID NO:101)

RHDDEELKEVVGQKKSFNIIQVVIGEKTNPPIQQAVALYLSIELKENTKEARADKL
 RKRFEADGGKNKFLVILDDWQFVDLEDIGLSPLPNKGVNFKVLLSRDSHVCTL
 MGAEANSILNIKVLKDVEGQSLFRQFAKNAGDDDDPAFNGIADSIASRCQGLPIAI
 10 KTIALSLKGRSKSAWDVALSRLENHKIGSEEVREVFKISYDNLQDEVTKSIFLLCAL
 FPEDFDIPTIELVRYGWGLKLFIEAKTIAREARNRLNTTERLRETNLLFGSDDIGCVK
 MHDVVVRDFVLHIFSEVQHASIVHGNVSEWLEENHSIYSCKRISLTCKGMSQFPKDL
 KFPNLSILKLMHGDKSLSFPENFYGKMEKVQVISYDKLMLYPLLSSLECSTNVRLH
 15 LHYCSLRMFDCCSIGNLLNMEVLSFANSIEWLPSTIGNLKKLRLLDLNCKGLRID
 NGVLKNLVKLEELYMGVNRPYQAVSLTDE

RG2H polynucleotide sequence (SEQ ID NO:102)

TGAAGGAGGTTGTGGAACGAAAGAAAATGTCAGTATTATTGTTCAAGTG
 GTCATAGGAGAGAACAAACCTATTGCTATTCAAGCTGTAGCAGA
 20 TTACCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATA
 AGCTTCGTAATGGTCGAGGCCGATGGAGGAAAGAATAAGTCCCTTGTA
 ATACTTGACGATGTATGGCAGTTGTCGATCTGAAGATATTGGTTAAG
 TCCTCTGCCAATAAAGGTGTCAACTCAAGGTCTTGTGACGTCAAGAG
 ATTACACATGTTGCACTCTGATGGGAGCCGAAGCCAATTCAATTCTCAAT
 25 ATAAGTTAACAGCTGTAGAAGGACAAAGTTGTTCCGCCAGTTGC
 TAAATGCGGGTGATGATGACCTGGATCCTGCTTCAATAGGATAGCAG
 ATAGTATTGCAAGTAGATGTCAAGGTTGCCATTGCCATAAAACCATT
 GCCTTAAGCTTAAAGGTAGAAGCAAGCCTGGCTGGGACCATGCGCTTTC
 TCGTTGGAGAACATAAGATTGGTAGTGAAGAAGTTGCGTGAAGTT
 30 TTAATATTAGCTATGACAATCTCAAGATGAGATTACTAAATCTATT
 TTACTTGTGCTTATTCTGAAAGATTGATATTCTACTGAGGAGTT
 GATGAGGTATGGATGGGCTTGAAATTATTAGAAGCAAAACTATAA
 GAGAAGCAAGAACAGGCTAACACACTGCACTGAGCGGCTAGGGAGACA
 AATTGTTATTGGAAGCGATGACATTGGATGCGTCAAGATGCACGATGT
 35 GGTGCGTGATTTGTTGCATATATTCTCAGAAGTCCAGCACGCTCAA
 TTGTCAACCATGGTAACGTGTCAGAGTGGCTAGAGGAAATCATAGCATC
 TACTCTTGAAAAGAATTCTTACATGCAAGGTATGTCTGAGTTCC
 CAAAGACCTCAAATTCCAACCTTCAATTGAAACTTATGCATGGAG
 ATAAGTCGCTGAGCTTCCATTGAAAGATGGAAAAGGTT
 40 CAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCCCTCATCACT
 TGAATGCTCCACTAACGTTGAGTGCTCATCTCATTATTGTCATTAA
 GGATGTTGATTGCTTCAATTGTAATCTCTAACATGGAAGTGCTC

AGCTTGCTAATTCTAACATTGAATGGTACCATCTACAATTGGAAATT
GAAGAAGCTAAGGCTACTAGATTGACAAATTGTAAAGGTCTCGTATAG
ATAATGGTGTCTAAAAAATTGGTCAAACTTGAAGAGCTTATATGGGT
GTTAATCATCCGTATGGAC

5

RG2H deduced polypeptide sequence (SEQ ID NO:103)

KEVVERKKMFSIIVQVVIGEKTNPPIAQQAADVYLSIELKENTKEARADKLKWFEA
DGGKNKFLVILDDVWQFVDLEDIGLSPLPNKGVNFKVLLTSRDSHVCTLMGAEAN
SILNIKVLTAVEGQSLFRQFAKNAGDDDLDPAFNRIADSIAsrcQGLPIAIKTIASLK
10 GRSKPAWDHALSRLENHKIGSEEVREVFKISYDNLQDEITKSIFLLCALFPEDFDIP
TEELMRYGWGLKLFIEAKTIREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVVR
DFVLHIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEFPKDLKFPNLSI
LKLMHGDKSLSPENFYGKMEKVQVISYDKLYPPLLSSLECSTNVRVLHLHYCSL
15 RMFDCSSIGNLLNMEVLSFANSNIEWLPSTIGNLKKRLLDTNCKGLRIDNGVLKN
LVKLEELYMGVNHPYG

RG2I polynucleotide sequence (SEQ ID NO:104)

AAGAAGAGCTGAAGGAGGTTGTGGAACAAAAGAAAAGTTCAATATTATT
GTTCAAGTGGTCATAGGAGAGAAGACAAACCTATTGCTATTCAAGCAAGC
20 TGTAGCAGATTCCCTCTCTATAGAGCTGAAAGAAAACACTAAAGAAGCAA
GAGCTGATAAGCTCGTAAATGGTCGAGGCTGATGGAGGAAAGAATAAG
TTCCTCGTNACTTGACGATGTATGCCNGTTGTTGATCTGAAGATAT
TGGTTTAAGTCCTCATCCAAATAAAGGTGTCANCTCAAGGTCTGTTGA
CGTCAAGAGATTCACATGTTGCACTCTGATGGGAGCTGAAGCCAATTCA
25 ATTCTCAATATAAAAGTTAAAAGATGTAGAAGGAAAAGTTGTTCCG
CCAGTTGCTAAAATGCGGGTGATGATGACCTGGATCCTGCTTCATTG
GGATAGCAGATAGTATTGCAAGTAGATGTCAAGGTTGCCATTGCCATC
AAAACCATTGCCCTAACGCTAACAGGTAGAAGCAAGTCTGCATGGACGT
TGCACTTCTCGTCTGGAGAACTATAAGATTGGTAGTGAAGAAGTTGTGC
30 GTGAAGTTTTAAAATTAGCTATGACAATCTCAAGATGAGGTTACTAAA
TCTATTTTTACTTGTGCTTATTCTGAAGATTGATATTCTAC
TGAGGAGTTGGTGAGGTATGGTGGGGCTTGAAATTATTAGAAGCAA
AAACTATAAGAGAAGCAAGAAACAGGCTAACACACTGCACTGAGCGGCTT
AGGGAGACAAATTGTTATTGAAAGTGTGACATTGGATGCGTCAAGAT
35 GCACGATGTGGTGCCTGATTTGTTGCTATATTCTCAGAAGTCCAGC
ACGCTTCAATTGTCACCATGGTAATGTGTCAGAGTGGCTAGAGGAAAAT
CATAGCATCTACTCTGTAAAAGAATTCTATTAACATGCAAGGGTATGTC
TGAGTTCCAAAGACCTCAAATTCCAACCTTCAATTGAAACTTA
TGCATGGAGATAAGTCGCTGAGCTTCTGAAACTTTATGAAAGATG
40 GAAAAGGTTCAGGTAATATCATATGATAAATTGATGTATCCATTGCTTCC
CTCATCACTGAATGCTCCACCAACCTCGAGTGCTTCACTCTCATGAAT
GTTCATTAAGGATGTTGATTGCTCTCAATTGTAATCTTCTCAACATG

GAAGTGCTCAGCTTGCTAATTCTGGCATTGAATGGTTACCATCTACAAT
TGGAAATTGAAGAAGCTAACGGCTACTGGATCTGACAGATTGTGGAGGTC
TTCATATAGATAATGGCGTCTAAAAAATTGGTCAAACTTGAAGAGCTT
TATATGGGTGCTAATCGTCTGTTGGAAAGTGCCAT

5

RG2I deduced polypeptide sequence (SEQ ID NO:105)

EELKEVVEQKKTFNIVQVVIGEKTNPIAIQQAVADSLIELKENTKEARADKLKW
EADGGKNKFLVILDDVW?FVDLEDIGLSPHPNKGV?FKVLLTSRDSHVCLMGAEA
NSILNIKVLDVEGKSLFRQFAKNAGDDLDPAFIGIADSIASRCQGLPIAIKTIALSL
10 KGRSKSAWDVALSRLENHKIGSEEVVREVFKISYDNLQDEVTKSIFLLCALFPEDFDI
PTEELVRYGWGLKLIEAKTIAREARNRLNTCTERLRETNLLFGSDDIGCVKMHDVV
RDFVLHIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEFPKDLKFPNLS
15 ILKLMHGDKSLSPENFYGKMEKVQVISYDKLMYPLPSSLECSTNRLVHLHECSL
RMFDCCSIGNLLNMEVLSFANSGIEWLPSTIGNLKKRLLDLCGGLHIDNGVLKN
LVKLEELYMGANRLFGKCH

RG2J polynucleotide sequence (SEQ ID NO:106) and (SEQ ID NO:107)

ATGTCCGACCCAACAGGGATTGTTGGTGCCTATTATAACCCAATTGCTCA
AACGGCCTGGTCCCCTAACAGACCATGTAGGCTACATGATTCCCTGCA
20 GAA-AATATGTGAGGGACATGCAAATGAAAATGACAGAGTTAAATACCTCA
AGAACATCAGTCAGAGGAACACATTAGCCGGAACACAAGAAATCATCTTCA
GATTCCATCTCAAATTAGGATTGGTGGACCAAGTAGAAGGGATCAGAG
CGAATGTTGCAAACCTTCCAATTGATGTCATCAGTTGTTGAGTCTCAGG
ATCAGGCACAAGCTTGGACAGAAAGCCTCAAGATAACTGAGCAGATCGA
25 AAGTCTAACGAGACAAAATCGCTGATTATCTGGACTGATGAACCTGTT
CCCTGGGAAGAGTTGGTCCATGATTGCACTCACCTCTGCAGCATCAAGT
GATCATCATGATGTCCTCCCTCAAGAGAGCAAATTTAGGAAAGCACT
AGAACGACTTGAACCCGTCAAAATCCCACATAATAGCCTATGGGGGA
TGGCGGGAGTGGGAAGACCACGATGATGAAGAAGCTGAAAGAGGTCTG
30 GAACAAAAGAAAACGTGCAATATTATTGTTCAAGTGGTCATAGGAGAGAA
GACAAACCTATTGCTATCCAGCAAGCTGTAGCAGATTACCTCTATAAG
AGCTGAAAGAAAACACTAAAGAAGCAAGAGCTGATAAGCTTGTAAACCGG
TTCGAAGCCGATGGAGGAAAGAATAAGTCCCTGTAATACTTGACGATGT
ATGGCAGTTTCGATCTTGAAGATATTGGTTAAGTCCTCTGCCAAATA
35 AAGGTGTCAACTCAAGGTCTGTTGACGTCAAGAGATTCACATGTTGC
ACTCTGATGGGAGCTGAAGCCAATTCTATTCTCAATATAAAAGTTTAAA
AGATGTAGAAGGAAAAGTTGTTCCGCCAGTTGCTAAAATGCGGGTG
ATGATGACCTGGATCCTGCTTCATTGGATAGCAGATAGTATTGCAAGT
AGATGTCAAGGTTGCCATTGCCATCAAAACCATTGCCCTAAGTCTAA
40 AGGTAGAAGCAAGTCTGCATGGGACGTCGCACTTCTCGTCTGGAGAATC
ATAAGATTGGTAGTGAAGAAGTTGCGTGAAGTTTAAAATTAGCTAT
GACAATCTCCAAGATGAGGTTACTAAATCTATTTTACTCTGTGCTT

ATTCCTGAAGATTGATATTCTATTGAGGAGTTGGTGAGGTATGGGT
GGGGCTTCAAATTATTTATAAGCAAAACTATAAGAGAAGCAAGAAC
AGGCTCAACAACACTGCACTGAGCGGCTTAGGGAGACAAATTGTTATTGG
AAGTCATGACTTGGGTGCGTCAAGATGCACGATGTGGTGCCTGATTTG
5 TTTGCATATGTTTCAGAAGTCAAGCATGCTCAATTGTCACCAGGT
AACATGTCAGAGTGGCCAGAGAAAAATGATACCAGCAACTCTGTAAAAG
AATTCTTAACATGCAAGGGTATGTCTAAGTTCTAAAGACATCAACT
ATCCAACCTTTGATTTGAAACCTATGCATGGAGATAAGTCGCTGTGC
TTTCCTGAAAACCTTATGAAAGATGGAAAAGGTTCAGGTAATATCATA
10 TGATAAATTGATGTATCCATTGCTTCCCTCATCACTGAAATGCTCCACTA
ACGTTGAGTGCTTCATCTCATTATTGTCATTAAGGATGTTGATTGC
TCTTCATGGTAATCTCTCAACATGGAAGTGCTCAGCTTGCTAATT
TAACATTGAATGGTTACCATCTACAATTGAAATTGAAAGAAGCTAAGGC
TACTAGATTGACAAATTGAAAGGCTTCGTATAGATAATGGTGTCTA
15 AAAAATTGGTCAAACCTGAAGAGCTTATATGGGTGTTAACCGTCCGTA
TGGACAGGCCGTTAGCTTGACAGATGAAAATGCAATGAAATGGTAGAAG
GTCCAAAAAAACTCTTGCACTAGAAATATGAGTTGTTAAATACAATGCT
CAAGTGAAGAATATATCCTCGAGAACATCTAACGATTCAAGATCTCAGT
GGGATGTTCTTACATGGATCTTCAGTAAAAGCAGGCACTCATACGAAA
20 ACACGTTGAAGTTGGCATTGACAAAGGCGAATATTGGAATCCGAATG
AACGGGTTGTTGAGAAAACGGAGGTTCTTGTAAAGTGTGGGGATAT
GTATCATCTTCAGATGTTAAGGTGAAGTCCTCTCGTTACAATTAA
GAGTCCTTGTGTTAGAGTGTGAGAGTTGAAACACCTCTCACACTT
GGTGTGCAAATACTTGTCAAAGCTTGAGCATCTAAAGTCTACAAATG
25 CGATAATATGGAAGAACTCATACATACCGGGGTAGTGAAGGAGATA
TTACATTCCCCAAGCTGAAGCTTATATTGATGGCTGCCAACCTA
TTGGGTTGTCCTAATGTCACGCAATTGAGCTACAAAATTGTGCA
AATGAAGCTTACAGCATTCCGGTTCAACAGCATTATCCGCGGAACA
AGTGGAAAGCATCTAGTTGAAAGAAGAGGTACATACATATAGTT
30 TATGTTAACATTTAAACAATCTTCAACTAAAAGTTCTAGAACATATA
TCTGTATTGATTGATGTGTTAGTGTGTTGGATGTGGCTATTAAAG
GATATTATTGGCAGGTTGTGATTCTAAGTTGGATATACTTGAAATT
ATGACATGGAGAATTAAAGGAAATATGGCCTAGTGAGCTTAGAGGT
GAGAAAGTTAACAGTTGAGAAAGATTAAAGTGAGAAATTGTGATAAAACTGT
35 GAATCTATTCCACACAATCCCAGTCTCTGCTGCATCATCTTGAAAGAGC
TTATAGTCAGAAATGTGGTCCATTGAAGAGTTGTCACACATCGACTTG
GATTGTGCCAGTGTAAATTGGAGAAGAAGACAACACAGCAGCTTAAGAAA
CATCAATGTGGAGAATTCAATGAAGCTAAGAGAGGTGTGGAGGATAAAAG
GTGCAGATAACTCTCGTCCCTCTTCGTGGCTTCAAGTTGTTGAAAAG
40 ATAATCATTACGAGATGTAAGAGGTTACAAATGTATTCACACCTATCAC
CACAAATTGATCTGGGGCACTTTGGAGATTTCAGTTGATTGTAGAG
GAAATGATGAATCAGACCAAGTAACCAAGAGCAAGAGCAGGTATGGATT
TCAATTACTCTTACTTAATTAATGATTAAGCCCCTGCTTTAATA

AAAAGGGGACAAACCATTCTGACTTAATGTTGCAATACAAGTCATGTA
 TAAGAGTGATTAACCTTTTTTATTATAAAAATACTACAAAACATGTTT
 TTTCATTATAGATCATGTATAAATGTTGACTAATTTTTCATGCCCTAAC
 TTTTGTGATAAATCATTAGAAATGTCACTAATTACTTTAGTATTAT
 5 AAAATAACTACAAAACATGTTTTTATTATAGATCATGTATATATCAC
 TAAAAAATATTATTCCCTTACACAAAAAAAAGGTTCAAGAAAGCCTGTA
 TTTCGAAATAACTAAAAGAAAATATTGATATTCACTAAGAGAAATTT
 10 TTCTAAACATGATCGAAATGATTAAAACCTAAATTAAAACAAAAAGA
 TTTTATATATGTTATNCAAAATTAAAATTGAAATTAAAGTTATAATT
 15 TNGTNTCACAAAGGGATATATAGAAAATATTATTTTGAGTCAT
 GCATAGTTGTATTTAAATGATTATTAACGTGGTAGGAGTGGAAACCA
 CTCAATCTAGTAGACCCACTATCACATGTCACATCAGCTTACATCTATT
 TTTCTTCTCCTTTTCATCTTTAAACTCATAACACNTAAAANTANC
 ATATTTCCAACACACTNAACTCATTGTCACATTATTATTTTAATTAA
 20 TTAATTNGAAAATTAAAANTAAANTAAANCNTAACATTTTAATTAAA
 AATATTAATCCAATAAAANTNCACGATAAATTAAAANGTTANTTG
 GAAAAAAANCC (SEQ ID NO:106)
 Sequence gap
 ATAACCCTTCAAGGGTCAACTCAAGTCCAAGTTAAAGTCAGGTCAAAA
 25 CCTTGGTAAAGTCAACTTGGTCAAAGTCACATCTACTGACTCACCT
 CACCGAGTTGGTCCACCAACTTGTGAGTCCCTTAATCCACAAACTCAA
 GAACCTCGATCCTACTCGTCAGTCTTCAGAAACTCTCGAGTTCCAT
 TACACAGAATGGGACCTTTGCTCATGACTCGCCAGTTCATCCTTGAA
 CTTGTCGAGTCTAGCTTACAGAGTTGAGTGTAGTCCTGACTCGT
 30 CGAGTTCTCCTGAACTCGTCAGTCCATCTCGTATAGTTGGGACATT
 GCCTTGAACTCACCGAGTTCATCATTGAACTCATCGAGTCCTCGATCTT
 CAAGTCCATAATCCTGTCCATCTGTTGAGTCCTCTTAGACTCAACCA
 GATTCCCTCAGAAACAGAAAAGGTTAGGGAACCAATTACCTGACTCGCGAG
 TCCCAAGAACGAATCCCCGAGTCCCCAATGTCCATGACCATAACATCGA
 35 TTTTCGTTGGGCTCATTGCATCCAAAGCATAGATCTAACCTCTAGGGTC
 CATATTACACGTAAAGCTACGAACCTGACGTCCATGCATGGGGGATTGG
 CTCAAATGGCATTAAAATGGGTTATCTGATGCATGGACTCCATGGC
 CATAAAGTTAACACCTTATGCCATGGGAATCTCAATGGTCCATATCT
 GAAGTTAACACTCTACAATATGTTCAAACCGAAGGTGGCTTAGAAATG
 40 CCCCAAAATGGCAAGATTCAAGCCTAAAGGAGATCTAACAAATGATAAG
 TCAAGGTTCAAGCTTTACCTGAAATAAGCTGGAAATGAAGCAAAATCT
 CTGGATCCACTTGTCTCTCAAGAACCCCCAAGCTCCACTTCTCTTC
 AAGTTCAAACAACTTTAAACACTCAAAAATGGCTCAAGAACACTCAAAA
 AGCTTAGGGTTCGAGTTAGGGCTTTGGAAAGCGAGAGGGACGATGGG
 GGCTGAAATGAGGCTAGAAAAAGTGTAAATAGGGGCAAACCCCTAAAT
 ATTAGGGTTCATCCAGGCAGCCCTACTCGTCAGTCGGGCTCCGACTC
 GTCGAGTAGGTCACTAAAACCCGCGTCCATAATCCAGTCACTCGACGA
 GTTGGGCCTCCAACTCGTCGATTCCGAGTGCAAAACGTTCAATTACTAA

ATTTAAATATGTACCAGGAACCGGGTGTACAGTTGAGACTTATACCTC
CATAAAGATAGATCTAGGTGCACATAGCCTGGATCCACAAGCTCCATGTCA
ACAAGCGACTCTTCAAGAAGTTCATTCTCCTTAAGCACCAAAAAAC
ACACAAAATCACCATGAAGCTCAAGAAATACTCAAATAGAGGATAGGGTT
5 TCGITCGTAGGGTTAGAGAGGGATGGAGGCTAGAGGAATGAGGGATAGAG
GCGAGTTAAGGTCTTAAATAGGGTCCAAGACCCCTAAATTAGGGTTAA
TCTGGCCAGACGAACGCAGGGTGTCCCAAATGCATATGTGTCCAATT
TCGTGTGCGCCATGCGTACCTCCCTGTACGCCATGTGTACCGGGTTGG
TCCAACCCCTCTAACCTCAAATGATCATAACTTGACCCCCCTATCTGTT
10 TTCGATGTTCTTATCCACGGAAAGGTAAACAAGAAGCCCTATACCTCT
ATAAACTTATTTAATCTGAAAACCAACCGAAATTAAATCCAAAATT
AAAAGTCCCAGACCAACACATTACCGATACCCCTGGGCTCCAAAACACA
AATTGAAAACCCGGATCATCCAAACTACATCATCCACCTCCAAATGAGCC
CAAACCTCAATTATTCAAGGGTTCTAACGCCTGTTAATGCCCACTCCTCGAT
15 TACCAACCCCGCAATGGGAAACGATTCAAAACAGGGCGTTACATAATTG
TGTGGTTTGTATTTTATTCCGGTGAAGGTGAAAGATCCAACACTATT
TTAATCTGTTGGCATTTCATCATTGCAACTGTTCTGAAAAAA
TACCTAAAATCAAATAACCATTTCAAATCCAAAATTATAAGAGAGAAT
TGTAAATGGACATGGAATCTTAAATCATTAACACAGTTCACTACACAAGT
20 TGCTAATTACATTCTGCTGTGCAGATTGAAATTCTATCAGAGAAAGAG
ACATTACAAGAACGCCACTGACAGTATTCTAATGTTGATTCCATCCTG
TCTCATGCACTCTTCATAACCTCCAGAAACTTATATTGAACAGAGTTA
AAGGAGTGGAGGTGGTGTGAGATAGAGAGTGGAGAGTCCAACAAGT
GAATTGGTAACAACCTACCCATAACCAACAACAGCCTGTTATATTCCAA
25 CCTCCAGCATTGGATCTAAGGGTATGGACAACATGATTGCGGTGG
AGTGCAGCAACTGGAATAAATTCTCACTCTTCCAAAACAACATCAGAA
TCCCCATTCCACAAACCTCACAACCATAAAATTGATTGATTGCAAGC
TAAGTACTTGTTCACCTCTCATGGCAGAACCTTCCAAACCTAAAGA
AAAGTCAATATAAAATGGTGTATGGTATTGAAGAAGTTGTTCAAACAGA
30 GATGATGAGGATGAAGAAATGACTACATTACATCTACCCACACAACCAC
CATCTGTTCCCTCATCTGATTCTCACTCTAAGTTCTGGAGAATC
TGAAGTGTATTGGTGGAGGTGGTGCAGGATGAGGGGAGCAATGAAATA
TCTTCATAATAACCAACTGCAACTACTGCTGTTCTGATCAATTGAGGT
ATGCTTGTTCATATTCAATTATTAAATTCTTTATTGCAA
35 TATTCTATAAATAACATTATACCCACTATAACTAAGATAATAATTAC
CTAGAGGGATGGATGCTATGACACAGCTGCTACACTTCAGAAACTCTAGT
AAGGGCAGTTATGGAAGTTCAATAAAATGATAATGGCATCTTGATGGG
TAATATAGGCAATTAAAGTTTATTCTGTTAAAGCAGTATTAGCAAGT
ACTGCCAGTAGGAGAGGAGAATATCACCTTGTGAAAATCTGGTCATT
40 GTACCCAGAATTAGTTAAATGTAACATTAGATATCAGGGGACATCAG
GTGACAGATATTGTAGAATAGAACAAATATAATTACCCAAA
TTTCTAAGGTTTCTGTTAAATATGTGCTTCTGATTGATT
TGCATTCTATATTAGGTGGTAAAGTGATTGCTCTCAATAAACCCC

GAA.ATTAATTAAAAAAAAACAAAAGTAAATTTGATATGGAGA
GCACTGGTATCATTAGTATATAAAAAACTAGATTTGAATTAAGTTTC
TTATATAAAAAGCTGTATATAGTTAATTAGTTTACATCATTTC
TGTGGTGTGCAGTTCTGAAGCAGGTGGTCTGGAGCTATGCC
5 AATACGCTAGAGAGATAAGTATAGAATTCTGCAATGCATTGTCAAGTGTG
ATTCCATGTTATGCAGCAGGACAAATGCAAAGCTCAAGTGCTGACAGT
CAGTTCTGTAATGGTCTGAAGGAGGTATTGAAACTCAATTAAGGAGGA
GCAGCAACAAAAACAACGAGAAGAGTGGTTGTGATGAAGGAAATGGTGG
ATTCCAAGAGTAAATAACAATGTTATTATGCTTCTGGTCTGAAGATATT
10 GGA.AATCAGCTTGTGGGGTTGGAACATATATTCACATTCTGCAC
TTGAAAGCCTGAGACAGCTCGAAGAGTTAACGATAATGAATTGCTGGTCA
ATG.AAAGTGATTGTGAAAGAAGGAAGAAGATGAATATGGAGAGCAGCAAAC
AACAAACAACGAAGGGACTTCTCTCTCTCTCTCTCTCT
CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
15 TTTCTTGTCTAAAGTCATTGTATTGGTCAATCTACCAGAGCTGGTAGG
ATTCTTCTGGGATGAATGAGTTCCGGTTGCCTTCATTAGATGAACCTTA
TCATCGAGAAATGCCAAAAATGATGGTGTACAGCTGGTGGTCCACA
GCTCCCCAACTCAAGTATATACACACAAGATTAGGCAAACATACTATTGA
TCA.AGAATCTGGCCTTAACCTTCATCAGGTATATATGTTCTTAATTGG
20 CATCATCTAATTAAAGAAAGATATCATTCTGCCAAGTAAATTACTTCAA
ACACATTCACACTGGTTCACTCTCAAGTTATGTTCTTAGGAAGGCCA
AAA TGGGAAAGCAAGATAGGGAAAAATAGTGTATTCACTGGAAAGGGTA
TTTAGGTATTTCTGTCAAAAGTTGTTATTGCAGGCTTTAGTACCTG
GAA.TCGTGTGGGAGGAGCATTATTCTGATTGCTGTTCTTAT
25 CATTCTCTTAGCCTCTCGAACAGCTAGAAACCCCTTTAATCTTTGAT
TTT.AAATGACAAAATTCTCCCTGTTACTCTATTGATTGTTCTTCA
TGGTTCTAAGTGAGTTATTGGCTCATCTGTTACTCTTTGATTGTTATT
TTC.ATAGCATGTTAGTCACTGAATCAAGCTTTCACTTCAACCAGGG
CAAAAGGTCAAAAGTAACCTACTTATGAGATCAAAACAGCAACCCATC
30 GGATAACTTTAGTTGGAGTTAATAGTTACAATTACCAATTGTGATTAATA
ATT.ATAATATCCTGTATTAAATTCAAAAAATTGGTACAGCACATATATGA
CATTCAAAAGGTTTGTTGACATATATATGCCCTGGCGTTCTTTA
TTGGACTTGCAGACCTCATTCAAAGTTATACGGTACACCTGGGCC
TGCTACTTCAGAAGGGACAACCTGGTCTTCTATAACTGATTGAATTAG
35 ATGTGAAATTAAATAAGGATGTTAAAAGATTATTCCATCCAGTGAGTTG
CTGCAACTGCAAAAGCTGGAAAAGATAAAATATAACAGTTGTGTTGGGGT
AG.AGGAGGTATTGAAACTGCATTGGAAGCAGCAGGGAGAAATGGAAATA
GTGGAATTGGTTTGATGAATCGTCACAAACAACACTACCAACTCTTGTC
AATCTTCAAACCTTAGAGAAATGAACCTATGGGTCTAGATTGTCTGAG
40 GT.ATATATGGAAGAGCAATCAGTGGACAGCATTGAGTTCCAAAACAA
CAAGAGTTGAAATTAGTAATTGCAACAGTTAGAACATGTATTACTAGT
TCATGGTTGGTAGTCTATCGCAACTCCAAGAGCTACATATAAGTCAGTG
CAAAACTTATGGAGGAGGTATTGTTAAGGATGCAGATGTTCTGTAGAAG

AAGACAAAGAGAAAGAATCTGATGGCAAGATGAATAAGGAGATACTTGCG
 TTACCTAGTCTAAAGTCCCTGAAATTAGAAAGCTTACCATCTCTGAGGG
 GTTAGCTGGGAAGGAGGATTTCATTCCCATTATGGATACTTAA
 GAATTGAGGAATGCCAGCAATAACCACCTCACCAAGGGAAATTCCGCT
 5 ACTCCACAACTAAGAGAAATAGAAACAAGATTGGCTCGGTTATGCAGG
 GGAAGACATCAAATCCTATTATAAGATCAAACAACAGGTAAATCAGA
 TCATTGTTGGTTAATAATTCTAAACTACATTGAAAAGTTCATGTAA
 GTTTTTATTATTGTCAAAAGCCGCAACCTATATTCAACTTTATATT
 ATGTACTTTATGCAGGATTCAAAAAAGCCAGGACTCTATTAAATGTGA
 10 AGTAAATACTAGAACAGGTAATTCTATTACATGTCTCCTGATTGCCTA
 TTAATTAAATGGCCTTCAGTTCATGGTTTGATGTATTCTCATGATG
 ACGTAATGTTAAATACCCCACTAGTTAATTGTTAGGTTGAATGTTGAT
 GACCAAAGGACTATATGTCGGAAGAATATTCAAGGAAAGAATTGTCAT
 CATATGAAGGCATTAATTAAAGAAGAACATGGATGCTATGAAGATGTTG
 15 GGAAAATATATGAATCAAATAACAAGCTACTCACTTATCTAAGTTGTTG
 GTTGAGGATGTTGATTAAATATTCAAATTCAATTGATGTTGATTATG
 GTTTATCAGTAGTGTATGGATAATGAGCAACTAACCTAAATTATG
 CTGTTGGTAAATGTTGACTCAAGTATGAAAATTAGGAATAACTTGTGA
 AAAATATATGCAAAAGTAGGATTGAGATTCAATGAAAAAAATTATGAA
 20 ACTATACTACTATAGTATATAAATAATTCAACTTACTGTTGGGTATATT
 GGAAGCACATATCATGAAAGTAACTAGAACGAGAATTGTTCCATCTC
 ATCTACTTATAGTTCCATTCTACTGTAAAAATCTGATTAAACTTTA
 GAGTTATTCTATTACCAACCAAATTTCATATAAAGGCCACAAG
 T (SEQ ID NO:107)

25

RG2J deduced polypeptide sequence (SEQ ID NO:108)

MSDPTGIVGAIINPIAQTLVPLTDHVGYMISCRKYVRDMQMCKMTELNTSRISAEEH
 ISRNTRNHLQIPSQIKDWLDQVEGIRANVANFPIDVISCCSLIRHKLQKAFKITEQI
 ESLTRQNLSIIWTDEPVPLGRVGSMIASTAASSDHHDVFPSREQIFRKALEALEPVQ
 30 KSHIIALWGMGGVGKTTMMKKLKEVVEQKKTCNIIVQVVIKEKTNPIAIQQAVADY
 LSIELKENTKEARADKLRKRFEADGGKNKFLVILDDVVQFFDLEDIGSPLPNKGV
 NFKVLLTSRDSHVCTLMGAEANSILNIKVLKDVEGKSLFRQFAKNAGDDDLDPFI
 GIADSIASRCQGLPIAIKTIALSKGRSKSAWDVALSRLENHKIGSEEVVREVFKISYD
 NLQDEVTKSIFLLCALFPEDFDIPIEELVRYGWGLKLFIEAKTIAREARNRLNNCTERL
 35 RETNLLFGSHDFGCVKMHDVVRDFVLHMFSEVKHASIVNHGNMSEWPEKNDTSN
 SCKRISLTCKGMSKFPKDINYPNLLILKLMHGDKSLCPENFYGKMEKVQVISYDKL
 MYPLLPSSLECSTNVRLVHLHYCSLRMFDCSSIGNLLNMEVLSFANSNIEWLPSTIG
 NLKKLRLLDLTNCKGLRIDNGVLKNLVKLELYMGVNRPYGQA VSLTDENCNEM
 VEGSKKLLALEYELFKYNAQVKNISFENLKRKVISVGCSLHGSFSKSRHSYENTLKL
 40 AIDKGELLESRMNGLFEKTEVLCLSVDVVKVSSSFYNLRVLVVSECAEL
 KHLFTLGVANTLSKLEHLKVYKCDNMEELIHTGGSEGDTITFPKLKLLYLHGLPNL
 LGLCLNVNAIELPKLVQMKLYSIPGFTSIYPRNKLEASSLLKEEVVIPEELIVEKCGSI
 EELFNIDLDCASVIGEEDNNSSLRNINVENSMLREVWRIGADNSRPLFRGFQVVE

KIIITRCKRFTNVFTPITTNFDLGALLEISVDCRGNDESDQSNQEQQIEILSEKETLQE
 ATDSISVVFPSCLMHSFHNLQKLILNRVKGVVVFEIESPTSRELVTTHHNQQP
 VIFPNLQHLDLRGMNDMIRVWKCSNWNKFFLPKQQSESFPHNLTINIDFCRSIKY
 LFSPLMAELLSNLKKVNIKWCYGIEEVVSNRDDEEMTTFTSTHTTILFPHLDSDL
 5 TLSFLENLKCIGGGAKDEGSNEISFNNTTATTAVLDQFELSEAGGVWSLCQYAR
 EISIEFCNALSSVIPCYAAGQMQLQVLTVSSCNGLKEVFETQLRRSSKNNEKSGC
 DEGNGGIPRVNNNVIMSLGLKILEISFCGGLEHIFTFSALESLRQLEELTIMNCWSMK
 VIVKKEEDEYGEQQTTTTKGTSSSSSSSSSSSSSSSPPSSSSKKVVVFPCCLKSIVLVNLP
 ELVGFFLGMNEFRPLSDELIEKCPKMMVFTAGGSTAPQLKYITRLGKHTIDQES
 10 GLNFHQDIYMPLAFSLLDLQTFSQSLYGDTLGPATSEGTTWSFHNLIELDVKFNKD
 VKIIIPSSELLQLQKLEKININSCVGVEEVFETALEAGRNGNSGIGFDESSQTTL
 VNLPNLREMNWLGLDCLRYIWKSQNWTAFEFPKLTRVEISNCNSLEHVFTSSMVGS
 LSQQLQELHISQCKLMEEVIVKDADVSVEEDKEKESDGKMNKEILALPSLKSLSKLES
 PSLEGFSLGKEDFSFPLLDTLRIECPAITFTKGNSATPQLREIETRFGSVYAGEDIKS
 15 SIIKIKQQDFKKAQDSI.CEVNTR

RG2K polynucleotide sequence (SEQ ID NO:109) and (SEQ ID NO:110)

TGGGATTCCATATATAAAAACATATATTTTATAAAGTGGGATTCCATTG
 TTTATATAGATTTTATTCACCAATAGACAATAGATTAAGATATA
 20 AAAACATGTCGGCTTTGACTAAAAATATAGATTTATGAATAGAATAT
 TCAATTGCTTAACCGTTAAAAAAATGAAAAGATGTCGATATAAAA
 TCTCATATGGGCCTTCTTACCATCAAATAGTAAAATAGTAAAAGATAC
 TTGTTTGGGCATGAACGTGACCATACTCAAAACCCATACAAAATCAAACGA
 ATCCCACATGGATGATGACGGATGGGTCGAGTAAATGTGTTTGGCCT
 25 TTTTTTCGAGAGAACAGAACAGCTTCTGCTCTCATCTCTTAGATTTG
 GGGATTTCTGGTTTCAGGGTTTGAGTGGAAACTAAATTGAAGCAA
 AAAGTATGGTATAATTGGTTCTAGTGAAATTGATGCTTCTATTACTAT
 CATCTTAAAATTGTCAAAACATTATGTATTAAATTATGAGATCGAAAGT
 GGTCTATGGCCAAGGTAATACAAGCTTACTCAATGAAATGAATCTAGG
 30 ATGCATCATGCATGTATTGGTTAGATTAAAGATTTCATCAAATTCTT
 TATCAAATTGTTGTATACCATGTTATGTAGGTGCTACCACAAGCCATAAC
 ATCGAGCAATGGAGTGTATTACTGGCATCTTAGCAACCCGTTGCTCAG
 TGTCTCATCGCTCCTGTGAAAGAACACCTTGCCTCTGATTTCATAC
 ACAATATGTAGGGATATGCTTACTGCAATGACGGAGTTGAATGCTGAA
 35 AAGACATTGTTGAAGAGCGGAAGAACAAACGTAGAAAATGTTTGAG
 GTTCCAAACCATGTCAACCGTTGGTGGAAAGATGTTCAAACAATCAACAG
 AAAAGTGGAACGTGTTCTAACGATAATTGCAATTGGTTCAATCTATGTA
 ATAGGTACATGCTCGCAGTGAAGCCTGGAGATAACTCAGGAGATCGAT
 CATGCCATGAAACAACCTCTCGGATAGAATGGACTGATGATTCAAGTG
 40 TTTGGGAAGAAATGATTCCACAAAGGCATCCACCTCTACACCATCAAGTG
 ATTACAATGACTTCGAGTCAAGAGAACACACTTTAGGAAAGCACITGAA
 GCACCTGGATCCAACCACACATCCCACATGGTAGCCTATGGGGATGGG

TGGAGTTGGGAAGACCACGATGATGAAGAGGGCTGAAAAATATTATAAAG
AAAAGAGGACGTTCATTATATTGTTGGTGGTATAAAGGAAAATATG
GATCTCATTCATCCAGGATGCTGTAGCAGATTATCTGGATATGAAGCT
AACAGAAAGCAATGAATCAGAAAGAGCCGATAAAACTCGTGAAGGGTTC
5 AGGCCAAATCAGATGGAGGTAAAGAATAGGTTCCCTCATATAACTGGATGAT
GTATGGCAATCTGTTAATATGGAAGATATTGGTTAAGTCCTTTCCGAA
TCAAGGTGTCGACTTCAGGTCTGTTGACCTCGGAAAACAAAGATGTT
GTGCAAAAATGGGAGGTGAAGCTAATTAAATTTCGACGTGAAATTCTTA
ACAGAAGAAGAAGCACAAAGTTGTTATCAATTGAAAAGTTCTGA
10 TACCCACCTTGATAAGATTGAAAAGCTATTGTAAGAAACTGTGGTGGTC
TACCCATTGCCATCAAAACCATAGCCAATACTCTTAAAAATAGAAACAAG
GATGTATGGAAGGATGCACTTCTCGTATAGAGCATCATGACATTGAGAC
AATTGCACATGTTGTTTCAAATGAGCTACGACAATCTCCAAAACGAAG
AAGCTCAATCCATTGGCTTGATTGTTGAGACTTTGAT
15 ATTCCTACTGAGGAATTGGTGAGGTATGGATGGGATTGAGAGTATTAA
TGGAGTGTACTATAGGAGAAGCAAGACACAGGTTGAACGCCTACATCG
AGCTGCTCAAGGATTCTAATTATTGATTGAAAGTGTGATGTTCACTGC
ATCAAGATGCATGATTAGTCGTGCTTGGATACGTTAATAG
ATTCAAGCATTCTTGATTGTTAACATGGTAATGGTGGTATGTTAGGGT
20 GGCCTGAAAATGATATGAGTGCTCATCTTGCAAAAGAATTCTTAAATA
TGC.AAGGGCATGTCGATTTCTAGAGACGTAAGTTCAAATCTCTT
GATTTGAAACTTATGCATGCAGATAAGTCTTGAAGTTCTCAAGACT
TTATGGAGAAATGAAGAAGCTTCAGGTTATATCATAACGATCACATGAAG
TATCCCTGCTTCAACATCACCTCAATGCTCCACCAACCTCGTGTGCT
25 TCA.TCTTCATCAATGCTCATTGATGTTGATTGCTCTTCTATTGGAAATC
TGGTGAATCTGGAAGTGCTCAGCTTGCTAATTCTGGTATTGAGTGGTTG
CCTTCCACAATCGGAAATTGAAGGAGCTAAGGGTACTAGATTGACAAA
TTGTGATGGCTTCTCGTATAGATAATGGTGTCTAAAGAAATTGGTGAAC
TTG.AAGAGCTTATATGAGAGTTGGTGGTCGATATCAAAGGCCATTAGC
30 TTC.ACTGATGAAAATGCAATGAAATGGCAGAGCGTCAAAAATCTTC
TGC.ATTAGAATTGAGTTCTCAAAAACAATGCTCAACCAAAGAATATGT
CATTGAGAATCTTGAACGATTCAAGATCTCAGTGGGATGTTTTAAG
GGAGATTCGGAAGATCTTCACTCTTGAACACAGTTGCGGTTGGT
CACCAACAGAACTGAAGTTCTGAATCTAGGCTTAATGAGTTGTTGAGA
35 AACACAGATGTTCTTATTAAAGTGTGGGAGATATGAATGATCTTGAAGAT
GTTGAGGTAAGTGGCACATCTCCTAAATCCTCTTCCACAATT
AAGAGTCCTTATCATTCTGAGTGTAGAGTTGAGATACCTTTCACAC
TTG.ATGTTGCAAACACTTGTCAAAGCTTGAGCATCTCAAGTTACGAA
TGCAGATAATGGAAGAAATCATAACACAGAGGGTAGAGGAGAAGTGAC
40 AATTACATTCCAAAAGCTGAAGTTTATCATTGTTGGCTACCAAATC
TGGTGGGTTGTGTTGAGATCTGACATAATTAAATCTACCAACTCACA
GAGTTGAAACTTAATGGCATTCCAGGTTCACAAGCATATATCCTGAAA
AGATGTTGAAACATCTAGTTGTTGAATAAGAGGTAAATGTGTTTATG

TTAATACAATAACATTTCAATTAAACCGTTCAAAATATATTGTATGA
TTTATTTGTTGGATGGGGTTATTAAATGGGTGATTATTCTCAGGTTG
TAATTCTAATTGGAGAAACTTGATATTAGTATATGAAGGATTGAAA
GAGATATGCCCTGTGAATTAGGGATGAGTCAGGAAGTTGATTTCTAC
5 GTTGAGAGTGATTAAGTAAGCAGTTGTGATAATCTTGTGAATCTATTCC
CGTGCAATCCTATGCCATTGATACATCACCTGAAGAGCTCAAGTGATA
TTTGTGGITCCATTGAAGTGTATTCAACATTGAGTTGGATTCTATTGG
TCAAATTGGAGAAGGCATCAACAATAGCAGCTTGAGAATCATCCAATTGC
AGAACTTAGGGAAGCTAAGTGAGGTGTGGAGGATAAAAGGTGCGGATAAC
10 TCTAGTCTCTCATCAGTGGTTCAAGGTGTGAAAGCATTATCGTTAA
CAAATGCAAGATGTTAGAAATGTATTACACCTACCACCAATTGG
ATCTGGGGGCACTTATGGAGATTGGATACAAGATTGTGGAGAAAAGAGG
AGAAACAAACGAATTGGTAGAGAGTAGCCAAGAGCAAGAGCAGGTATGGCT
TTCAATTCTACTTCTTAATGAAGGATTAAGCTCTGCTTTGAA
15 TAAAAAAGTGGATGAATGACTAAATTGGGAATGCCACCCGGAAAGTTATC
AACCATTTAGCTACACCATTGGTAACATGTTGCAATAATGCATAA
TATAATTAAAAAATGGTCATTGATAAAATGTAACCAACCTTTTATTAA
TTAAAATGCTACAATAAATGATTTCTTTATTATATCATTAAAC
AATAAGCTTAAAGATGTTAAATAGCCAATGTCAGTTAGATCGTAAC
20 AATTTTTATTAACTAGTTTAGTTAAGATATCACTCATTATTATTTA
TAGAAAAAAAGACAAGATTGGCTAATCCTCATAAGAATTGGAAGATTAA
GCAAAATATAGAGCTTCCAAACATAGCCAATAGTTCTTGCAGGTC
CCATCTACGAAATTATCAATAGATTGCGATTGGCACCCGGGA
AATTCCATTAAATTAAAAAAAGTTCAAGCCATTGTAGTTGGCACCTG
25 CAAAATGGTAGTTGCACCTGCGGAATCACCTTCACCATTGCGATCT
ATGACTTGTGAAAATGTTAATTGTGAAATGGTCATGTGCACCTCATGAG
AAATCGAAATGGTCAGTAATATGACTTTTATATAAATATGATGGTGG
CATATATTATAGGAAAATATAGCTGCACGATATTAAATTAGTGAAT
TAGTTAAGTACGATAAGTACACAAATTATATGATGAAGTAC
30 TCAATTAGGACGACTCGGGCAATGAAATCATCATTAAATAGGAGCAATG
AAATCATTTCGAAAATGTTACAAATGAATAAAATATTAAATTAAACT
TAAAACATTGTTAGTAGTTGAAATTACAAACTGAAATTGTTGTAT
TTATTAAACATTATAATGTTACTATGATTTCCTGTTGCAAAT
ATTCCCTAAAAATCCACCTAAAATCAAATAATTACATTCAAGTTG
35 AAAAATGAAATCGTATGATATAACCGTGTGGATGTGGAAATTATATA
CAGTTACTAATTACATTGTTGGATATGTCGCGAGATTGATATT
GCAATCCCATTCACTCTCACACACTCTTCCAAAACCTCCGTAAACTG
TTGGAAAAGTATGAAGGAGTGGAGGTGGTTGAGATAGAGAGTCAA
CAAGTAGAGAATTGATAACAATTACCCATAATCAACAACCAACTTCCC
40 AACCTTGAGTTAGGATATAAGTTATGGACAGCATGAGTCATGTATG
GAAGTGCAACTGGAATAAATTCTCATTCTCAAAAACAACAGTCAGAAT
CCCCATTCTGTAATCTCACACCATACATTCAATATTGCCAAAGCATT
AAGTACTGTTTCAACTCTCATGGAAAACCTCTTCCAACCTAAAGAA

GGTCGAGGTAAGAGAGTGTATGGTATTGAAGAAGTTGTTCGAACAGAG
ATGATGAAGATGAGGAAAAGACTACATTTACATCTACATCTCTGAAAAAA
AGCACTAATTGTTCCCTCGTCTTGAATCTCTCGCTTTATCAACTTCC
AAATCTCAAGTGTATTGGTGGTGGTCTGCCAACAGTGGGAACAATG
AAATATCTCTTGATAATTCCACTACTACTACTCTTTGTTGATCAATCT
AAGGTATGTTTTTTNTNCCCTT (SEQ ID NO:109)
Sequence gap
CCTCCCTAATAATACATGTTATGCACACTATACTAACATATTAGACACGT
AAAGGATAAAATGCTATGCCTCATATAATACGTTATTTATAATCTTAA
10 ACAATCAAATTATTAAACAAATAACTAAGTGTGAGCAAAGGCAGGTACC
CGACTAAATTGCCAAAACCAGTCTGGTGGTCTGGGAATGTTGGGCCAG
GTCGTTAAAACGTCTACACACCGTTCTTAAATCACAGATCCGCTCTC
ATACTGTGAACCCGGTTTAATTAAAAGAAAATTTCATTATAAAGTAA
ATGACTTAAACCATTACAAACAAACAAAATTACCATACAAATGTTGGAC
15 TATCATTATTGCAACATAAAACTGAAAATACACATATTCTCTGATA
TCAGCATGAGTGGCTGGCTAACCCAAAATCCATGCATTGTAGATG
TGTGTTACAACACATAGTATCAATGAAAGGCATATTAGGCTAGAATT
TAACAATCTGAATAATATTCCCTAAAACATAATCATCATCAACCAACT
AATATAAAACCATTGGGTTCGTCATTAGGTACAAACATAGATTTC
20 TAAGCTTGTGTATTAAACATATGCTTCTAAACTTAATTGATTGCA
TTCCAAAATTAGGTGAAAGTGGTATGTCATTGTTGCTTTCAAC
ATTAATTGTACAAAACCAAAACTACATAATTGATGTAGATATCATAACA
ATTGTGTTATTAGTATATAAAAACAAATTGAAATTGAATTCTTATA
CAAAGTTGTGTCTATGTATACATGTTATGTAGGTAAAGACAATTAGT
25 CTCTGTTAAGTATGGAGTTAATTAGACTAATTTCATGTGTTG
CAGTTTATCAGGCAGGTGGCTTTGGACGTTATGCCAATACTCCAG
AGAGATAAAATATAAGGGAGTGTATGCATTGCAAGTGTAAATTCCATGTT
ATGCAGCAGGACAGATGCAAAATGTTCAAGTGTGAATATACAGGTGC
AACTCAATGAAGGAGTTATTGAAACTCAAGGGATGAACAACAACAATGG
30 TGACAGTGGTTGTGATGAAGGAAATGGTTGATACCAAGCAATTCCAAGAC
TAAATAACGTTATTGCTACCAATCTAAAGATATTGAAGATTGAAGAT
TGTGGTCATCTGGAACATGTATTCACATTCTGCACTTGGAAAGCCTGAG
ACAGCTCGAAGAGTTAACGATAGAGAAATGCAAGGCAATGAAAGTGTAG
TGAAGGAAGAAGATGAATATGGAGAGCAAACAACAAGGCATCTCGAAG
35 GAGGTTGTTGCTTCCCTCGTCTCAAGTCCATTGAACTGGAAAATCTACA
AGAGCTCATGGTTCTACTTAGGGAAGAATGAGATTGAGATTGAGCTGGCCTTCAT
TGGATAAGGTTATGATCAAGAATTGCCAGAAATGATGGTGGTTCACCT
GGTAGTCCACAGTCCCAAGCGCAAGTATATAAATACAAGCTTGGCAT
ATATGGGATGGAGGAGGTACTTGAAACTCAAGGGATGAACAACAATAATG
40 ATGACAATTGTTGTGATGATGGAAATGGTGGAAATTCCAAGACTAAATAAC
GTTATTATGTTCCAATATAAGATATTGCAAATCAGCAATTGTGGCAG
TTTGGAACATATTCACATTCTGCACTTGAAAGCCTGATGCAGCTCA
AAGAGTTAACATAGCGGATTGCAAGGCAATGAAAGTATTGTGAAGGAG

GAATATGATGTAGAGCAAACAAGGGTATTGAAGGCTGTGGTATTTCTTG
 TCTAAAGTCATTACACTATGCCATCTACCAGAGTTGGTGGGTTCTTCT
 TGGGGAAAATGAGTTCTGGTGGCCTTCATTGGATAAGGTTACCATCATT
 GATTGCCACAAATGATGGGTTCACACCTGGTGGTCAACAACTCCCA
 5 CCTCAAGTACATACACTCAAGCTTAGGCAAACATACTCTTAATGTGGCC
 TTAATTCAAGTCACAACACTGCATATCATCAGGTATAATTATTCT
 TTNACACCATCTAATTATGGAATCATGACGCTAATTACAGTATTAAACAC
 (SEQ ID NO:110)

10 RG2K deduced polypeptide sequence (SEQ ID NO:111)

MECITGIFSNPFAQCLIAPVKEHLCLLIFYTQYVGDMLTAMTELNAAKDIVEERK
 NQNVEKCFCVPNVNWRLEDVQTINRKVERVLNDNCNWPNLCNRYMLAVKAL
 EITQEIDHAMKQLSRIEWDDSVPLGRNDSTKASTSTPSSDYNDFESREHTFRKAL
 EALGSNHTSHMVALWGMGGVGKTTMMKRLKNIKEKRTFHYIVLVVIKENMDL
 15 ISIQDAVADYLDMKLTESNESERADKLREGFQAKSDGGKNRFLIILDDVVQSVN
 MEDIGLSPFPNQGVDFKVLLTSENKDVCAMGVEANLIFDVKFLTEEEAQSLFY
 QFKVVSDFTHLDKIGKAIRNCGLPIAKTIANTLKNRNKDVWKDALSRHHED
 IETIAHVVFQMSYDNLQNEEAQSIFLLCGLFPEDFDIPEELVRYGWGLRVFNGV
 YTIGEARHRLNAIYIELLKDSNLIESDDVHCIMHDLVRAFVLDTFNRFKHSLIV
 20 NHGNGGMLGWPENDMSASSCKRISLICKGMSDFPRDVFKFPNLLILKLMHADKS
 LKFPQDFYGEVKLQLVISYDHMKYPLLPTSPQCSTNLRLVHLHQCSLMFDCCSI
 GNLLNLEVLSFANSGIEWLPSTIGNLKELRVLDLTNCGLRIDNGVKKLVKLEELY
 MRVGGRYQKAISFTDENCNEMAERSKNLSALEFEFFKNNAQPKNMSFENLERFKIS
 VGCYFKGDFGKIFHSFENTLRLVTNRTEVLESRLNELFEKTDVLYLSVGDMNDLED
 25 VEVKLAHLPKSSSFHNLRLVLIISECIELRYLFTLDVANTLSKLEHLQVYECDNMEEII
 HTEGRGEVITFPKLKFLSLCGLPNLLGLCGNVHIINLPQLTELKNGIPGFTSIYPEK
 DVETSSLLNEVVIPNLEKLDISYMKDLKEIWPCELGMSQEVDVSTLRVIKVSSCDN
 LVNLFP CNCMPPLIHHEELQVIFCGSIEVLFNIELDSIGQIGEGINNSSLRIQLQNLGK
 LSEVWRIKGADNSSLLISGFQGVESIIVNKCKMFRNVFTPTTNFDLGALMÈRIQDC
 30 GEKRRNNELVESSQEQQ

RG2L polynucleotide sequence (SEQ ID NO:112)

GGAAGACACAATGATGCAAAGACTGAAGAAGGTTGCCAAGAAAATAGAA
 TGTTCAAGTACATGGTCGAGGCAGTTATAGGGAAAAGACAGACCCAATT
 35 GCTATTCAACAAGCTGTAGCCGATTACCTCGTATACAGTTCAAAGAAAG
 CACTAAACCAGCAAGAGCTGATAAGCTCGTAATATTGATGACGCTGGCAG
 CTGNAGACGGAAGAATAAGTTCTCGTAATATTGATGACGCTGGCAG
 TCCGTTGATCTGGAAAGATATTGGNTTAAGTCCTTTCCAATCAAGGTGT
 CGACTTCAAGGTCTTGTGACTTCACGAGACGAACACGTTGCACAATGA
 40 TGGGGGTTGAAGCTAATTCAAGTTATTAAATGTGGACTTCTAACTGAAGTA
 GAAGCACAAAGTCTGTTCCAGCAATTGTAGAAACTTTGAGCCCGAGCT
 CTGTAAGATAGGAGAAGTTATCGTAAGAAAGTGTGCGGTCTACCTATTG

CCATCAAAACCATGGCGTGTACTCTAAGAAATAAAAGAAAGGATGCATGG
 AAGGATGCACTTCACGTATAGAGCACTATGACATTGCTAGTGTTGCGCC
 TAAAGTCTTGAACAAAGCTATCACAATCTCCAAGACAGGGAGACTAAAT
 CCGTGTGTTGATGTGTGGTTGTTCTGAAGACTTCAATATTCTTACCC
 5 GAGGAGTTGATGAGGTATGGATGGGCTTAAAGCTATTGACAGAGTTA
 TACAATTAGAGAAGCAAGAACCGGCTAACACACCTGCATTGAGCGACTTG
 TGCAGACAAATTGTAATTGAAAGTATGATGATGTTGGGTGTCAAGATG
 CATGATCTGGTGCCTGCTTTGTTGGGTATGTATTCTGAAGTCGAGCA
 TGCTTCATTGTAACCATGGTAATATGCATGGGTGGACTAAAATGATA
 10 TGAACGACTCTGCAAAACAGTTCTTAACATGCGAGAGTGTGTCTGAG
 TTTCCAGGAGACCTCAAGTTCCAACCTAAAGCTTTGAAACTTATGCA
 TGGAGATAAGATGCTAAGGTTCTCAAGACTTTATGAAGGAATGGAAA
 AGCTCCAGGTAATATCATAACCATAAAATGAAGTATCCATTGCTTCCCTCG
 TCACCTCAATGCTCACCAACCTCGAGTGCTTCATCTCATCGGTGTT
 15 ATTACGGATGCTTGTATTGCTCTGTATCGAAATTGACGAATCTGGAAG
 TGTTGAGCTCGCTAATTCTGGCATTGAACGGATACCTCAGCAATCGGA
 AATTGAAAGCTTAGGCAACTGATCTGAGAGGTCGTTATGGTCTTG
 TATAGAACAGGGTGTCTGAAAAATTGGTCGAACTGAAGAACATTATA
 TTGGAAATGCATCTGCGTTAGAGATTATAACTGCAATGAGATGGCAG
 20

RG2L deduced polypeptide sequence (SEQ ID NO:113)

EDTMMQRLKKVAKENRMFSYMVEAVIGEKTDPIAIQQAVADYLRIQFESTKPAR
 ADKLREWFKAHS?DGKNKFLVIFDDVVQSVLEDIGLSPFPNQGVDFKVLLTSRDE
 HVCTMMGVEANSVINVGLLTEVEAQSLFQQFVETFEPELCKIGEVIVRKCCGLPIAI
 25 KTMACTLRNKRKDAWKDALSRIEHYDIRSVAPKFETSYHNLQDRETKSVFLMCG
 LFPEDFNIPIEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQTNLLIESDDVGC
 VKMHDLVRAFVLGMYSEVEHASIVNHGNMHGWTKNDMNDSCKTVSLETCEVSEF
 PGDLKFPNLKLLKLMHGDKMLRFSQDFYEGMEKLQVISYHKMKYPLLPSSPQCST
 NLRVLHLHRCSLRMLDCSCIGNLTNEVLSFANSGIERIPSAGNLKKLRQLDLRGR
 30 YGLCIEQGVNLVELEELYIGNASAFRDYN岑NEMA

RG2M polynucleotide sequence (SEQ ID NO:114)

GGGGAAGACACAATAGATGCAAAGGCTGAAGAAGTTGCCAAAGAAAAGAG
 AATGTTCAGTTATATCATTGAGGCGTTATAGGGAAAAGACAGACCCCCA
 35 TTTCCATTCAAGGAAGCTATATCATATTACCTTGGTAGAGCTCAATGCA
 AATACTAACAGTCAGTAAGAGCTGATATGCTTCGTCAAGGGTTCAAGGCCAA
 ATCTGATGTAGGTAAGGATAAAATTCTAATAATAACTCGACGATGTATGGC
 AGTCTGTTATTGGAAGATATTGGATTAAGTCCATTCAAATCAAGGT
 GTTAACTTCAAGGTCCGTAAACATCACGAGACCGACATATTGCACTGT
 40 GATGGGGGTTGAAGGTCAATTGATTTAATGTGGACTTCTCACAGAAG
 CAGAATCAAAAAGATTGTTCTGGCAGTTGTAGAAGTTCTGATCCTGAG
 CTCCATAAGATAGGAGAAGATATTGTAAGTAAGTGTGTGGTCTACCCAT

TGCCATTAAAACCATGGCATGTACACTTAGAGATAAAAGTACGGATGCAT
 GGAAGGATGCACTGTCCTCGTTAGAGCATCATGACATTGAAAATGTTGCC
 TCTAAAGTTTAGAGCGAGCTATGACCCTCCAAGACGAGGAGACTAA
 ATCCACTTTCTATGTGGATTGTTCCAGAAGATTCCAATATTCTA
 5 TGGAGGAGTTGGTGAGGTATGGGTGGGGATTGAAATTATTTAAAAAGTG
 TATACCATAAGAGAACGAACTAGGCTCAACACTTGCATTGAGCGGCT
 CATCTATACCAATTGTTGATAAAAGTTGATGATGTTCAAGTCATCAAGA
 TGCACTGATCTCATCCGTTCTTGGATATGTTCTAAAGTTGAG
 10 CATGCTTCGATTGTCACCATGGTAATACGCTAGAGTGGCCTGCAGATNA
 TNTGCACGACTCTGTAAAGGGCTTCATTAACATGCAAGGGTANATGTG
 AGTTTGTGGAGACCTNAANTTCCAACCCTAATGATTAAACTTATG
 CATGGAGATAATCGCTAAGGTTT

RG2M deduced polypeptide sequence (SEQ ID NO:115)

15 GEDTIDAKAEEVAKERMF SYII EAVIGEKTD PISI QEAISY YLGVELNANTK SVRAD
 MLRQGFKA KSDVGKDKFLI LDDVWQSVDLEDIGLSPFPNQGVNF KVLLTSRDRHI
 CTVMGVEGH SIFNVGLL TEAESKRLFWQFVEGSDPELHKIGEDIVSKCCGLPIA IKT
 MACTLRDKSTD AWDKDALSRLEHHDIE NVASKVFRASYDHLQDEETKSTFFLCGLFP
 20 ED SNIPMEELV RYWG GLKL FKKV YTIREARTRLNT CIERLI YTNLLIKVDDVQCIKM
 HDLIRSFV LDMSK VEHASIVNHGNTLEWPAD??HDSCKGLSLTCKG?CEFCGDL?F
 PTLMILKLMHGDKSLRF

RG2N polynucleotide sequence (SEQ ID NO:116)

25 AGGTAAAATCCATAACCCTAAATGTTGGTACGCTCATATATCAAATTGCG
 TGTTTGTGAATGAAAAAAGCATGCTCAAAAACCAGTGTAAAGGCACGG
 TATATGACATATTATAGTTACTGATAACAAATTATGATAATTGGGTT
 TACRTAAGTTAGGATT CGTACTTCAACCAATGTAATAGTTTGAGT
 CTATCTATGTATTTGGGAATCACATTAGCAACGGGATTGTACTAGTAAT
 TCGAAAAAGTCTTAAATAATTCTGTTATAATTATGAATAGTT
 30 TAGCGACATCTAATATTAAGAATGTATCTGATATTGAATTAATGTCC
 TTAATGTGAACATAGACCTTCCATTACTAATGCCTAATTATTAGTT
 CTAATCAATAAATTAAATTCTGTTATGCTCTAAGACAATAAAAT
 CCATGATTACCTTAAATATTAACAAAAATGACCATAAAATAAAAAA
 ATTAGGATACCAACCCCCCGCCATGCCAATGTCTAAATATTCTTGAT
 35 GCTTTGCTTTCCCTCTTCCCTGTTAGTCTATTATTCTGGAGAGTT
 GAGAGAGTTCATACAAGAAAATTCAAGAAGAAAGCAAAGGTCCAGGT
 TTCTCTTCTTAATTATGTATTAACCTACAAGCATTACCGATCC
 ATGGTTTTGTGTATGTTCAAAATTGAAACTAGATTGGACTTTGC
 CCTTGATGATT CATAAGATATTGCATGGAGTTGAGATTGTGAAGAAAAG
 40 TGGTGAATAGAAAGAGCAAGTGAATCCAGATATAGTATTGGTAATATG
 ATGATGAGATAGAGATATGTTAAACTGGCTAGAAAATTGTTAATTG
 AAATTAGGTGTTGAATTGAAAGATAACCAAGCTAATAACTAATTAGTT

ATGCTAAWTAGTTATAAAGAACAAACACTCTAGTTTTTTTTCATGA
TTTCAACCTCTTGACCAAACAAATTATAGCAAAATTGAATATCATT
CTCTGCAATCAATCTAACCTTGTATTATCATCATGTCTAAAATTGCC
ACAAGTTATTTCAAAGTCATATTGGATTATGAAAGGACTATTTTACC
5 AATTACATCTTACTTTATGGGCCAAGCTAATACAATCCGACTAAACTA
AAGGAATATGGGATGCATATAGTTGCTTCCCATTATAGATTCTATCT
AATTTGTCTATTGTAACATTAGGTGCCACCACAAGTAAATTGTTAAA
TGGATATCGTTAATGCCATTCTAACCAAGITGTCGAGACTCTCATGGTA
CCCGTTAAGAAACACATAGGGTACCTCATTCTGCAGGCAATATATGAG
10 GGAAATGGGTATCAAATGAGGGGATTGAATGCTACTAGACTTGGTGTG
AAGAGCATGTGAACCGGAACATAAGCAACCAAGCTGAGGTTCCAGCCAA
GGCAGGGGTTGGTATGAAGAAGTAGGAAAGATCAATGCAAAAGTGGAAA
TTTCCTAGCGATGTTGGCAGTTGTTCAATCTAACGTTAGACACGGGG
TCGGAAAGAGAGCCTCCAAGATAATTGAGGACATCGACAGTGTCA
15 GAACACTCTATCATCTGGAATGATCATTCCATTCTCTAGGAAGAAT
TGATTCCACGAAAGCATCCACCTCAATACCATCAACCGATCATGATG
AGTTCCAGTCAGAGAGCAAACCTTCACAGAACGACTAAACGCACTCGAT
CCTAACCAACAAATCCCACATGATAGCCTATGGGAATGGCGGAGTGGG
GAAGACGACAATGATGCATCGGCTAAAAAGGTTGTGAAAGAAAAGAAA
20 TGTAAATTATTGTTGAGGCGGTGTAGGGAAAAAACAGACCCCCATT
GCTATTCAATCAGCTGTGGCAGATTACCTAGGTATAGAGCTCAATGAAA
AACTAAACCAGCAAGAACTGAGAAGCTCGTAAATGGTTGTGGACAATT
CTGCTGGTAAGAAGATCCTAGTCATACTCGACGATGTATGGCAGTTGTA
GATCTGAATGATATTGGTTAACGCTTACCAAATCAAGGTGTCGACTT
25 CAAGGTGTTGACATCACGAGACAAAGATGTTGCACTGAGATGGGAG
CTGAAGTTAATTCAACTTTAATGTGAAAATGTTAACAGAACAGAAC
CAAAGTTATTCCACCAATTGTAGAAATTCCGGATGATGTTGATCGTGA
GCTCCATAATATAGGAGTGAATTGTAAGGAAGTGTGGCGGTCTACCCA
TTGTCATCAAACCATGGCGTGTACTCTTAGAGGAAAAAGCAAGGATGCA
30 TGGAAAGAATGCACTCTCGTTAGTGAACTACAACATTGAAAATATAGT
GAATGGAGTTTAAAATGAGTTACGACAATCTCAAGATGAGGAGACTA
AATCCACCTTTGCTTGTGGAATGTTCCGAAGACTTAAATATTCT
ACCGAGGAGTTGGTGGAGGTATGGATGGGGGTGAAATTATTAAAAAGT
GTACTATAGGAGAACAGAACATCAGGCTAACACATGCATTGAGCGGC
35 TCATTACACAAATTGTTGATTGAAGTTGATGATGTTAGGTGCATCAAG
ATGCATGATCTGCGTGTGTTGGATATGTATTCTAAAGTCGA
GCATGCTCCATTGTCACCATGGTAATACACTAGAGTGGCATGTGGATA
ATATGCACAACCTCTGAAAAGACTTCAATTACATGCAAGGGTATGTCT
AAGTTCTACAGACCTCAAGTTCCAAACCTCTGATTTGAAACTTAT
40 GCATGAAGATATATCATTGAGGTTCCAAAACCTTATGAAGAAATGG
AGAAGCTTGAGGTTATATCCTATGATAAAATGAAATATCCATTGCTTCCC
TCATCACCGCAATGCTCCGTCAACCTTGCCTGTTCATCTCCATAAATG
CTCGTTAGTGTGATGTTGACTGCTCTGTATTGAAATCTGTCGAATCTAG

AAGTGCTTAGCTTGCTGATTCTGCCATTGACCTGTTGCCCTCCACAATC
GGAATTGAAAGAAGCTAACGGCTACTGGATTGACAAATTGTTATGGTCT
TTGTATAGCTAATGGTGTCTTAAAAAATTGGCAAAACTGAAAGAGCTCT
ATATGACAGTGGTAATGGAGGAGTCGAAAGGCATCAGCCTCACTGAG
5 GATAACTGCAATGAGATGGCAGAACGTTCAAAGACCTTCTGCATTAGA
ACTTGAGTTCTTAAAACAATGCTCAGCCAAGAATATGTCATTGAGA
AGCTACAACGATTCCAGATCTCAGTGGGTGCTATTATATGGAGCTCC
10 ATAAAGAGCAGGCACTCGTATGAAAACACATTGAAGTTGGTTATTGACAA
AGGTGAATTATTGAATCTGAATGAACGGCCTGTTAAGAAAACAGAGG
TGTTATGTTAAGTGTGGAGATATGAATGATCTGAAGAATRTTGAGGTT
15 AAGTCATCCTCACAACYTCTCAATCTTCTCGTTCAACAATTAAAGAGT
CCTTGTGTTCAAAGTGTGCAGAGTTGAAACACTTCTCACACCTGGTG
TTGCAAACACTTAAAAAAGCTTGAGCATCTGAAGTTACAAATGTGAT
AATATGGAAGAACTCATACGTAGCAGGGTAGTGAAGAAGAGACGATTAC
ATTCCCCAAGCTGAAGTTTATCTTGTGGCTACCAAAGCTATCGG
20 GTTGTGCGATAATGTCaaaATAATTGAGCTACCACAACTCATGGAGTTG
GAACCTGACGACATCCAGGTTACAAGCATATATCCATGAAAAAGTT
TGAAACATTAGTTGTTGAAGGAAGAGGTAATATAAATTAAATGCT
AATACATTACAAAGGATCTTCAGTAAATCTTCAAAATATATTGAA
25 TTTGATTGTATGGGTATTATTGTGGATGGACTATTAATAATGATTA
TCTTGCAGGTTCTGATTCTAAGTTAGAGAAACTGCATGTTAGTAGTATG
TGGAAATCTGAAGGAGATATGGCCTTGCATAATTAAATATGAGTGAGGAAGT
TAAGTTCAGAGAGATTAAAGTGAAGTAACGTGATAAGCTGTGAATTGTT
TTCCGCACAAGCCCATACTCTGCTCGTCATCTGAAGAGCTAAAGTC
30 AAGAATTGTGGTTCAATTGAATCGTTATTCAACATCCATTGGATTGTGC
TGGTGCAACTGGAGATGAATACAACAACAGTGGTGAAGAATTATTAAAG
TGATCAGTTGTGATAAGCTGTGAATCTCTTCCACACAATCCATGTCT
ATACTGCATCATCTGAAGAGCTTGAAGTCGAGAATTGTGGTTCCATTGA
ATCGTTATTCAACATTGACTGGATTGTGCTGGTGAATTGGCAAGAAG
35 ACAACAGAAGCAGCTAACAGAAACATCAAAGTGGAGAATTAGGGAAGCTA
AGAGAGGTGTGGAGGATAAAAGGTGGAGATAACTCTCGTCCCCCTGTTCA
TGGCTTCAATCTGTTGAAGCATAAGGGTACAAATGTAAGAGGTTA
GAAATGTATTACACCTACCACCAAAATTAAATCTGGGGCACTTTG
GAGATTCAATAGATGACTCGGGAGAAAACAGGGAAAATGACGAATCGGA
40 AGAGAGTAGCCATGAGCAAGAGCAGGTAAAGGATTCAATTCACTTCKT
ACTTAATTAAATGATTAAGCTCCTGCTTTTRAATAAAAAGGGACAAACC
ATTTCATGACTTAATGTAGCAATACAAGTCATGTATAAGAGTGACCAACT
CTTTTTATTATAAAATGACTACAAATATTTCATTAGAGATCA
TGTATAAAATGTGACTAATTTCATCACCTAACCTTGTGATAAAATCTT
TATAAATGTCACTAGTTACTTTCAGTAAATAACAAATTAAATAAATTAA
TCAACAAAAAGCATCAACTAAAAAAATCCCACAACCCGTAATAATTAAA
ATAAAAGGATTAAACATCTAACGAAACAATTCTAAACATGATT
TGGACCAAATATCACCAGCAACTCAAGTTGGAATCGATTAGCTAAAAA

CTTGACCACATAATTAGATAGATGAGAGTTGAAGCTAAAGTGCCTATAT
AAGTCGTTCATCTTTCTTGATCTGATAGCAAGTGAATSATTT
CTTCTCAAAATTGATAAAAATCTACATTATAAAGAGACTAGCTGAAAA
5 AAAATGGCTAGGTGGTCTTGGTAGATGAAGATGGAAGGGAGA
GTAGATTCAAAGACACAAACACATCTCATTATTATTATTATTATT
TTATTATTTGATATCTGCTCATATTGTTACAGATATGTGAGGTCT
ATTAAATCTTTAAATATATAAAAATAACATAATGAGAAAATTAA
ATAAAAGAATAAATTAAATAAGGGACAATAGCTTTTGGTAAGACAAGG
10 ACCAAAAGCGAACAAAAGTAAACAGTAGGGACCATCCGATTTAAAAAAT
TAATTAGGGACCAAAACATAAATTCCCCAAACCATAGGGACCATCGT
GTAATTACTCTGCTTTGTTCATATTGGTAACTATTTT
TTGTACATATCTAGGTAACGAACCTGTTGAAAGTGTTCACATCTACGATG
TGACCTACTACAACCGATCATAATGGTCATATGAACACTCCAACAAG
15 TTTGTTATCTAGGTGTGTACAAAAAAACGATAGTTACCATGATGTGAACA
TACCAAAAAATTAAATTACCTAGCAAGTTATTCCCATTAGGTGTAT
GGAAACAGITCCGTGAGACCGTGACTGGATGGTAGATAATTAGTAAA
CTTAACCCCTCAATTAAACCTACCTTTCTTATTAACCTAATTCAAGCT
AAATTCTGATTCTTGTGAAAGTAAGTTGCATCTTATGTTGTATTAT
20 CTTGTTGCATAGGATCCTAGCATCTTAATAATTATTGAAGGTGAA
AGATCCAACATATTAAATCTGTTGGCATTTCATCATTGCAACTGTT
TCTGAAAAAAA::TACCTAAAATCAAATAACCATTTCATATCCAAA
TTATAAGAGAGAATTGTTAACGGACATGGAATCATAATCATTAAACACAG
TTCAGTACACAGGTTGCTAATTACATTCTTGTGCAGATTGAAATT
TATCAGAGAAAGAGACATTACAAGAAGCCACTGGCAGTATTCAAATATT
25 GTATTCCCATTCTGTCTCATGCACTCTTCAAACTCCATAAAACTTAA
CTTGAACAGAGTTGAAGGAGTGGAGGTGGTGTGAGATAGAGAGTGAGA
GTCCAACAAGTAGAGAATTGTAACAACACTACCATAACCAACAACCT
ATTATACTTCCACCTCCAGGAATTGATTCTATGGAATATGGACAACAT
GAGTCATGTGTGAAAGTGCAGCACTGGAATAAATTCTCACTCTCAA
30 AAGAACAAATCAGAATCCCCATTCCACAAACCTCAGTAACACATACATATTAT
GAATGCAAAAGCATTAAGTACTTGTTCACCTCTCATGGCAGAACTTCT
TTCCAACCTAAAGCATATCGAGATAAGAGAGTGTGATGGTATTGAAGAAG
TTGTTCAAAAGAGATGGTGGAGGATGAAGACATGACTACATCTAC:::::
35 :::GCACACAAACCAACCACTTTCCCTCATCTTGTGATTCTCACTCTAA
GCAACTGAAGAATCTGAAGTGTATTGGTGGAGGTGGCCAAGGGATGAGG
GGAGCAATGAAATATCTTCAATAATACCACTGCAACTACTGCTGTTCT
GATCAATTGAGGTATGCTTGTACATATTCAATTATTATTAAATTCC
TTGTTAATTCTTCTTGTCAATTCTATGAAAAAAATCACCAA
40 TCACAAATAAGAGATTAAACTTTATTCAACACCCATGCGGACTCAAGA
ATGGGATTGGAGGCATATAAAGTTACATTCAATTGAAACAAGTATTACCA
TTTATTGTTATTATCATTTCATATCATTTACTGATAACATTCTT
TTACTTTCTAATTAGAAAAGGTCCACATGCTAATTAGGTTCCATT
TATGTGAATCCTCTATTCTGTCTGTAATCAAGCATCTAGATTATTATC

CATTTTCATAATTGTGTTATATTGACAGTTTTCTTTTATAGTTGT
 AATTGCAACCTGTCATATWTTMWWKKCWWWATKYWMWWARTAATACATTT
 TATAACCWCTATACTAAAGATA

5 RG2N deduced polypeptide sequence (SEQ ID NO:117)

LGKTTMMHRLKKVVKEKKMFNFIVEAVVGEKTDPIAQSAVADYLGIELNEKTKPA
 RTEKLRKWFDNSAGKKILVILDDVWQFVDLNDIGSPLPNQGVDFKVLLTSRDKD
 VCTEMGAEVNSTFNVKMLIETEAQSLFHQFVEISDDVDRELHNIGVNIVRKCGGLPI
 VIKTMACTLRGKSKDAWKNALLRLVNNNIENIVNGVFKMSYDNLQDEETKSTFLL
 10 CGMFPEDFNIPTEELVRYGWGLKLFKKVYTIGEARIRLNTCIERLIHTNLLIEVDDVR
 CIKMHDLVRAFVLDMSKVEHASIVNHGNTLEWHVDNMHNCKRLSLTCKGMSK
 FPTDLKFPNLSILKLMHEDISLRFPKNFYEEMEKLEVISYDKMKYPLPSSPQCSVNL
 CVFHLHKCSLVMFDCSCIGNLSNLEVLSFADSайдLLPSTIGILKKLRLLDLTNCYGL
 CIANGVFKKLVKLEELYMTVVNGGVRKAISL

15

RG2O polynucleotide sequence (SEQ ID NO:118)

TTGTAAAACGACGGCCAGTCGAATCGTAACCGTTGTACGAGAACCGCTG
 TCCTCTCCTTCATTTGAATCATGATATTGAATATCGATACTTTGACTG
 TAGCTTTGGTCGATTTTAGCAAGATACTGGCAAACCCATT
 20 GGCTATTTAGCCCCAAATATGAAATGGACTGGATTGTTTTTCCTTTC
 TAACACGCACACATCTGGCGATCAGTATCACTCCATTATGAAGACCTAGT
 CAAATTCAATTACGTTCACTCGTTCTCAAAGTTCAAAGTTCCAACCTT
 CCAACTTCCCTCTTTTTCTTCCTCGATTCTGATTGAATCCGAT
 TCTCGACGAAGGAGAGCTGGTCAGAGGGCTGTGATTCTGAGTCTGA
 25 CCTCCGAATCTAGCTGGATTATTTCGACACACCAGACCGTACAGGT
 TGCTCATCCCGAAATACTGCTTGCAAACCTGTTGATCATCGCTAGGAA
 ATTAAAGTTCTTTGGCTCTGTTACTGAATCAGTAGCTTGCAACTTG
 CTCATTATAAGCTGATCCATATTTACATATCTTGAAGAATAATAGGT
 ACTGACTTTACCTTCTGATGAGAGCGATTAAGAGATAACCTCTGAAAA
 30 TCCATTGGTGAAGGGATCTGGTTAGTTTAAAGGATTGCTACAAC
 AGTATCCCACAAACGATCTATTCCCATTNACTCATCCGCTCAAGATCT
 ATCCACCTTATATATGTTAATTGGAGTCTCCATGGTCAATGAATCT
 AGGATGCATTAGAACGCCAATCCATTACAAGTTCATCCAATTTCATG
 TGACAAGTTGGTTACTATGTAGGTACTTCCACAATTAAAGAATTCCA
 35 GCAATGGATGTTGTTAATGCCATTCTAAACCAGTTGCCAGACACTTAT
 GGAACCTGTTAAGAAACATCTAGGCTACATCATTCCAGCACAAACATG
 TGAGGGATATGAGTAACAAATGAGGGAGTTGAACGCTGCAAGACATGCT
 GAAGAAGACCACTGGACAGGAACATAAGAACCTCGTCTGAGATTCAA
 TCAAGTTAGGAGTTGGTTAGAAGAAGTAGAAAAGATCGATGCAAAAGTAA
 40 AAGCCCTCCTAGTGTACCGCTTGTGCAAGTCAGTCTCAAGATCAAACAT
 GAAGTCGGAAGGGAAAGCCTGAAAGCTAATTGAGATTGAAAGTGCAC
 AAGACAAACACTTTGATCACCTGGACTGATCATCCCATTCTGGAA

AAGTTGATTCCATGAAGGCATCGATGTCCACAGCATCAACCGATTACAAT
GACTTTCAGTCAGAGAAAAAAACTTTACTCAAGCATTGAAAGCACTTGA
ACCAAACAACGCTTCCCACATGATAGCGTTATGTGGATGGGTGGAGTGG
GGAAGACCACAATGATGCAAAGACTAAAAAAGTTGCTAACAAAATAGA
5 ATGTTCAGTTATATGGTTGAGGCAGTTATAGGGGAAAAGACGGACCCAAT
TGCTATTCAACAAGCTGTAGCGGATTACCTTCGTATAGAGTTAAAAGAAA
GCACTAAACCAGCAAGAGCTGATAAGCTTCGTGAATGTTCAAGGCCAAC
TCTGGAGAAGGTAAGAATAAATTCTTGTAAATACTTGATGACGTCTGGCA
GTCTGTTGATCTAGAAGATATTGGTTAAGTCCTTCCAATCAAGGTG
10 TCGACTTCAGGTCTTATTGACTTCACGAGACGAACATGTTGCACAGTA
ATGGGAGTTGGATCTAATTCAATTCTTAATGTGGACTCTAATAGAAGC
AGAAGCACAAGTTGTTCCAACAATTGTAGAAACTCTGAGCCCGAGC
TCCATAAGATAGGAGAAGATATTGTAAGGAAGTGTGCGGTCTACCTATT
GCCATCAAAAACCATGGCATGTTACTCTAGAAATAAAGAAAGGATGCTT
15 GAAGGATGCACTTCGGTATAGAGCACTATGACCTTCGCAATGTTGCGC
CTAAAGTCTTGAACAGAGCTACCACAATCTCATGACAAAGAGACTAAA
TCAGTGTGTTGATGTGTTGTTCCGGAAAGACTTCATATTCTAC
TGAGGAGTTGATGAGGTATGGATGGGATTAAAGATATTGATAGAGTCT
ATACATTATAGAAGCAAGAACAGGATCAACACCTGCATTGAGCGACTG
20 GTGCAGACAAATTGTTAATTGAAAGTGTGATGTTGGGTGTCAAGAT
GCA TGATCTGGTCCGTGTTTGTGTTAGGTATGTATTCTGAAGTAGAGC
ATGCTTCAGTTGTCAACCATGGTAATATACCTGGATGGACTGAAAATGAT
CCGACTGACTCTTGTAAAGCAATTCTATTAAACATGCGAGAGTATGTCTGG
AAACATTCCAGGAGACTCAAGTTCCAACCTAACGATTTGAAACTTA
25 TGCATGGAGATAAGTCGCTAACAGATTCCACAAGACTTTATGAAGGAATG
GAAAAGCTCCAGGTTATATCATACTGATAAAATGAAGTATCCAATGCTTCC
CTTGTCTCCTCAATGCTCCACCAACCTTCGAGTGCTCATCTCCATGAAT
GTTCATTAAAGATGTTGATTGCTCTGTATTGAAATATGGCGAATGTG
GAAGTGTGAGCTTGTAAATTCTGGCATTGAAATGTTACCTTCCACTAT
30 CGGAAATTAAAGAAGCTAACGTTACTTGATTAACAGATTGTCATGGTC
TTCATATAACACACGGTGTCTTAACAATTGGTCAAACATTGAAGAGTTG
TATATGGGATTTCTGATCGACCTGATCAAACCTGTGTTAATATTAGCAT
GACAGATGTCAGCTACAATGAATTAGCAGAACGTTAAAAGGCCCTTCTG
CATTAGAGTTCCAGTTGAAACAAATGCCAACAAATAATATGTCG
35 TTTGGGAAACTTAAACGATTCAAGATCTCAATGGGATGCACTTATATGG
AGGATCAGATTACTTAAAGAAAACGTATGCTGTCCAAAACACATTGAAGT
TGGTTACTAACAAAGGTGAACATTGGACTCTAGAATGAACGAGTTGTT
GTTGAAACAGAAATGCTTGTAAAGTGTGATGATATGAATGATCTTGG
TGA TGTTGTGTGAAGTCCTCACGTTCTCCTCAACCTCTGTGTTCAAAA
40 TTCTAAGAGTCTTGTGTTCCAAGTGTGTTGAGTGAAGATACCTTTC
ACAATTGGTGTAGCCAAGGATTGTCAAATCTGAGCATCTGAAGTTGA
TTCATGTAATAATGGAACAACTCATATGTATTGAGAATGCTGGAAAAG
AGACAATTACATTCTAAAGCTGAAGATTATCTTGAGTGGCTACCA

AAGCTTCGGGTTGTGCCAAAATGTCAACAAACTGAGCTACCACA
 CATAGAGTTGAAACTTAAGGCATTCCAGGGTTCACATGCATTATCCGC
 AAAACAAGTGGAAACATCTAGTTGTAAGGAAGAGGTAGATATATGT
 TTTATGTTAACACAAGTAAAAAATCTTTAACAAAAAGTTCACTATA
 5 TATATCTATATGTCTATAATTGATTATATGATGATTAGTGGATG
 TGGCTATTAGGGATGATTATTGCAAGGTGTGATTCCAAGTGGAGA
 CACTCAAATTGATGAGATGGAGAATTAAAGGAAATATGGCATTATAAA
 GTTAGTAATGGTGAGAGAGTTAACGTTAGAGAAAGATTGAAGTGAGTA
 TGATAAGCTTGTGAATCTATTCCACACAACCCATGTCCTGCTGCATC
 10 ATCTTGAAGAGCTTGAAGTCAAGAAATGTGGTCCATTGAATCGTTATT
 AACATCGACTTGGATTGTGATGCCATAGGAGAAGAACATGAG
 GAGCTTAAGAAACATTAAAGTGAAGAATTCACTGGAAGTTAAGAGAAGTGT
 GGTGTATAAAAGGTGAAATAACTCTGCCCCCTGTTCTGGCTTCAA
 GCTGTTGAAAGCATAAGCATGAAAGTTGAAGAGGTTAGAAATGTATT
 15 CACACCTACCACCAATTAAATATGGGGCACTTTGGAGATATCAA
 TAGATGACTGTGGAGAATACATGGAAAATGAAAATCGGAAAAGAGTAGC
 CAAGAGCAAGAGCAGGTATGGATTCAATTCACTTCTTACTTACTAA
 GGATTAAGCTTCTGTTTTGAATAAAAAGGGACATCTCTAATAATG
 CACATCTAAATTAAAAGTATTAAATTGTCATAGCAGCGTATAACAT
 20 CTTCTAATAATTATCTGAAGGTGAAAGATCCAACACTTCTAATTGTT
 AACAAATTCAATCATTGCAAATGTTCTTAAAAATTAAATTACCTGAAA
 TCAAAACAATCTCTCAAATCCAAAATTATGAGACAGAATTGAGAAGGG
 ATGTGAAATTATAAACCATTAACACAATTCCATGCTCACGTTACTAATT
 CATTCTGTTGGATATATATGTACAGACTGATATTGTCAGAGGAAG
 25 TGAAATTACAAGAAGTCACTGATACTATTCTAATGTTGATTACATCG
 TGTCTCATACACTCTTTATAACAACCTCCGTAACACTCAACTTGGAGAA
 GTATGGAGGAGTTGAGGTGTGTTGAGATAGAGAGTTCAACAAGTAGAG
 ATTGGTAACAACATACCATAAACAACAACAACAACACCTATATT
 CCCAACCTTGAGGAATTATCTATATTATGACACATGAGTCATGT
 30 ATGGAAGTGCACAAACTGGAATAAATTTCACAACAATCAGAACCCCCAT
 TCCACAAACCTCACACCACATGTCGATTGCAAAAGCATTAAGTAC
 TTGTTTCACCTCTCATGGCAGAACCTTCCACCTAAAGAGAAATCAA
 TATTGACGAGTGTGATGGTATTGAAGAAATTGTTCAAAAGAGATGATG
 TGGATGAAGAA

35

RG2O deduced polypeptide sequence (SEQ ID NO:119)

MDVNLKPVAVETLMEPVKKHLGYIISSTKHVRDMSNMRELNAARHAEEDHLD
 RNIRTRLEISNQVRSWLEEVEKIDAKVKALPSDVTAACSLKIKHEVGREALKLIVEIE
 SATRQHSLITWTDHPIPLGKVDSMKASMSTASTDYNDFQSREKTFTQALKALEPNN
 40 ASHMIALCGMGGVGKTTMMQLKKVAKQNRMFSYMVEAVIGEKTDPIAQQA
 DYLRIELKESTKPARADKLREWFKANSSEGKKNFLVILDDVVQSVLEDIGLSPFP
 NQGVDFKVLLTSRDEHVCTVMGVGSNSILNVGLLIEAEAQSLFQQFVETSEPELHKI

GEDIVRKCCGLPIAIKTMACTLRNKRDAWKDALSRIEHYDLRNVAPKVFETSYHN
 LHDKETKSFLMCGLFPEDFNIPEELMRYGWGLKIFDRVYTFIARNRINTCIERL
 VQTNLIESDDVGCVKMHDLVRAFVLGMYSEVEHASVNVHGNIPGWTENDPTDSC
 KAISLTCESMSGNIPGDFKFPNLTILKLMHGDKSLRFPQDFYEGMEKLQVISYDKMK
 5 YPMPLPLSPQCSTNLRLVHLHECSLKMFDSCSIGNMANVEVLSFANSGIEMPLSTIGN
 LKKLRLLDLTDCHGLHITHGVFNNLVKLEELYMGFSDRPDQTRGNISMVDVSYNE
 LAERSKGSALEFQFFENNAQPNNMSFGKLKRKFISMGCTLYGGSDYFKKTYAVQ
 NTLKLVTNKGEELDSRMNELFVETEMLCLSVDMMNDLGVCVKSSRSPQPSVFKIL
 RVFVVSKCVELRYLFTIGVAKDSLNEHLEVVDSCNNMEQLICIENAGKETITFLKLKI
 10 LSLSGLPKLSGLCQNWNKLELPQLIELKLKGIPGFTCIYPQNKLETSSLLKEEVVIPKL
 ETLQIDEMENLKEIWHYKVSNGERVKLRKIEVSNCDKLVNLFPHNPMSLLHHLEEL
 EVKKCGSIESLFNIDLDCVDAIGEEDNMRSRLNIKVNSWKLREVWCIGENNSCPL
 VSGFQAVESISIESCKRFRNVFTPTTNTNFNMGALEISIDDCGEYMEMENEKSEKSSSEQ
 EQTDILSEEVKLQEVTDTISNVVFTSCLIHSFYNNLRKLNLEYGGVEVVFEIESSTS
 15 RELVTTYHKQQQQQPIFPNLEELYLYMDNMSHVWKCNNWNKFLQQSESPFHNL
 LTTIHMSDCKSIKYLFSPLMAELLSNLKRINIDECDGI

RG2P polynucleotide sequence (SEQ ID NO:120)

CCCATTGCTATTCAAGGAAGCAGTAGCAGATTACCTCNGTATAGAGCTAA
 20 AGAAAAAAACTAAATCNGCAAGAGCTGATATGCTTCGAAAATGTTAGTTG
 CCAAGTCCGATGGTGGTAAAATAAGTTCTAGTAATACTTGACGATGTA
 TGGCAGTTGTTGATTAGAAGATATCGGTTAACGTGAGTGTAGATGTTGCA
 AGGTGTTAACTTCAAGGTCTGCTAACATCACGGGATGTAGATGTTGCA
 CTATGATGGAGTCGAAGCCAATTCAATTCTAACATGAAAATCTTACTA
 25 GATGAAGAAGCACAAAGTTGTTCATGGAGTTGTACAAATTGAGTGA
 TGTTGATCCAAGCTTCATAAGATAGGAGAAGATATTGTAAGAAAGTGT
 GTGGTTGCCTATTGCATAAAACCATGGCCCTACTCTTAGAAATAAA
 AGCAAGGATGCATGGAGTGTGACTTCTCGTTAGAGCATCATGACCT
 TCACAATTGTAATGAAGTTTGGATTAGCTACGACTATCTTCAAG
 30 ACCAGGAGACTAAATATCTTTGCTTGTGGATTGTTCCGAAGAC
 TACAATATTCCCTGAGGAGTTAATGAGGTATGGATGGGGCTAAATT
 ATTAAAAAAAGTGTATACTATAAGAGAAGCAAGAGCCAGACTCAACACCT
 GCATTGAGCGGCTTATCCATACCAATTGTTGATGGAAGGAGATGTTGTT
 GGGTGTGAAAGATGCATGATCTAGCACTTGCTTTGTTATGGATATGTT
 35 TTCTAAAGTCAGGATGCTCAATTGTCACCATGGTAGCATGTCAGGGT
 GGCTGAAAATGATGTGAGTGGCTTGCACAAAGAATTCAATTACATGC
 AAGGGTATGTCGGTTCCATAGACCTCAACTTCAAACCTCACAAT
 TTTAAAACCTATGCATGGAGATAAGTTCTCAAGTTCTCCAGACTTT
 ATGAACAAATGGAAAAGCTCAAGTTGATCGTTCATGAAATGAAATAT
 40 CCGTTCTCCCTCGTCTCCTCAATTGCTCCACCAACCTCGAGTTCT
 TCATCTCCATCAATGCTCATTGATGTTGATTGCTCTGTATTGGAAATC
 TGTTAATCTGGAAGTGTGAGCTTGCTAATTCTGGCATTGAATGGTTA

CCTTCCAGAATTGAAATTGAAGAACGTAAGGCTACTAGATTGACAGA
 TTGTTTGGTCTCGTATAGATAAGGGTGTCTAAAAAATTGGTCAAAC
 TTGAAGAGGTTATATGAGAGTTGCTGTCGAAGCAAAAAGCCGGAAAT
 AGAAAAGCCATTAGCTCACAGATGATACTGCAATGAGATGGCAGAGCG
 5 TTC

RG2P deduced polypeptide sequence (SEQ ID NO:121)

PIAQEAVADYL?IELKEKTKSARADMLRKMLVAKSDGGKNKFLVILDDWQFVDL
 EDIGLSPLPNQGVNFKVLLTSRDVDVCTMMGVEANSILNMKILLDEEAQSLFMEFV
 10 QISSDVPDKLHKIGEDIVRKCCGLPIAKTMALTLRNKSKDASDALSRLHHDLHN
 FVNEVFGISYDYLQDQETKYIFLLCGLFPEDYNIPPEELMRYGWGLNLFKKVYTIRE
 ARARLNTCIERLIHTNLLMEGDVVGCVKMHDLALAFVMDMFSKVQDASIVNHGS
 MSGWPENDVSGSCQRISLTCKGMGFPIDLNPNTILKLMHGDKFLKFPPDFYEQ
 MEKLQVVSFHEMKYPFLPSSPQYCSTNRLVLHLHQCSLMFDCSCSIGNLFNLEVLSF
 15 ANSGIEWLPSRIGNLKKLRLLDLTD CFG LRIDKGVLK NLVKLEEVYMRVA VR SKKA
 GNRKAISFTDDNCNEMAERS

RG2Q polynucleotide sequence (SEQ ID NO:122)

TGGGAAGACACAGT GATA GAAA AAAAAAGAAT GTTGTGGAAAAGAGGA
 20 AAATGTTGATTATGCTGTTGCGGGTATAGGGAAAAGACGGACCCT
 ATTGCTCTCAGAAA ACTGTTGCGGATTACTTGCAATTGAGCTAAATGA
 AAGCACTAAACTAGCAAGAGCAGATAAAACTTGCAAATGGTCAGGACA
 ACTCGGATGGAGGTAAAGAAAAAGTTCTCGTAATACTCGACGATGTTGG
 CAATCTGTGATTGGAAGATATTGGTTAAGTACTCCTTTCAAATCA
 25 AGGTGTCAACTCAAGGTTGTTGACATCAGAAAGAGAGAAATTGCA
 CAATGATGGGAGTTGAAGCTGATTTAATTCTCAATGTCAAAGTCTTAGAA
 GAAGAAGAACGACAAAAGTTGTTCCAGTTGAGAAATTGGTGACCA
 ATACCACGAGCTTCATCAGATAGGGTACATATAGTAAAGAAGTGTATG
 GTTACCCATTGCCATTAAAACCATGGCTCTTACTTAAGAAATAAAAGA
 30 AAGGATTCATGGAAGGACGCAC TCTCGTTAGAGGACCATGACACTGA
 AAAATGTTGCAAATGCAGTTTCGAGATGAAC TACCGCAATCTACAAGATG
 AGGAGACCAAAGCCATT TTTTGCTT GCGGTTGTTCCCCGAAGACTTT
 GATATT CCTACTGAGGAGTTGGTGAGGTATGGATGGGGCTTAAATCTATT
 TAAAAAAGTGTATACCATAAGAAAGGCAAGAACGAGATCGCATACATGTA
 35 TTGAGCGACTCTTGGATTCAAATTGTTGATTGAAAGTAACGATATT CGG
 TCGTCAAGATAACAGATCTGGTGC CGCTTTGTTGGATATGTATTG
 TAAAGTTGAGCATGCTCAATTGTCACCATGGTAATATGCGGACCGAAT
 ATAATATGGCTGACTCTTGCAAAACAATTCAATTACATACAAGAGTATG
 TCTGGGTTGAGTTCCAGGAGACCTCAAGTTCAAACCTAACAGTTT
 40 GAAACTTATGCANGGAGATAAGTCTCAAGGTTCTCAAGACTTATC
 AATCAATGGAAAAACTCAGGTTATATCATATGATAAAATGAAGTATCCA
 TTGCTTCCCTCATCACCTCAATGCTCCACTAACATCCGAGTGCTCGTCT

CCATGAATGTTCATTAAGGATGTTGATTGCTCTGTATTGGAAAGCTAT
 TGAATTGGAAGTCCTCAGCTTTTAATTCTAACATTGAATGGTACCT
 TCCACAATCAGAAATTAAAAAGCTAAGGCTACTAGATTGAGATATTG
 TGATCGTCTCGTATAGAACAAAGGTGCTTGAAAATTGGTCAAACTTG
 5 AAGAACTTTACGGATACATCAGCGTTACAGA

RG2Q deduced polypeptide sequence (SEQ ID NO:123)

GEDTVIEKKKNVVEKRKMFYAVVAVIGEKTDPIALQKTVADYLHIELNESTKLAR
 ADKLCKWFKDNDGGKKFLVILDDVWQSVLEDIGLSTPFPNQGVNFVKLLTSR
 10 KREICTMMGVVEADLILNVKVLEEEEAQQLFLQFVEIGDQYHELHQIGVHIVKKCYG
 LPIAIKTMALTLRNKRKDWSKDALSRLEDHDTENVANAVFEMNYRNLQDEETKAI
 FLLCGLFPEDFDIPEELVRYGWGLNFLKKVYTIRKARTRSHTCIERLLDSNLIESN
 DIRCVKIHDLVRAFLDMYCKVEHASIVNHGNMRTEYNMADSKTISLTYSKMSG
 15 FEFPGDLKFPNLTVLKLMDKSLRFPQDFYQSMEKLRVISYDKMKYPLLPSPPQCS
 TNIRVLRLHECSLRMFDCSCIGKLLNLEVLSSFFNSNIEWLPSTIRNLKKLRLLDLRYC
 DRLRIEQGVLKLNVKLEELYTGYTSATE

RG2S polynucleotide sequence (SEQ ID NO:124)

ATTTGGGGTTTACATITAATTTTGTGCATGAATGTGAAAATAGACTG
 20 CTTATTGATTCTTGTGTTCATGGAGTTGATTTCAATTACTACCTT
 ACAATTGCTCAGTGATAGATTCCATTAAATTGCTAATTGGTGCCTTC
 TAAATATGTAGGAGCTACTAAAGCAAAATATCGAGCAATGTCGGACCC
 AACGGGGATTGCTGGTGCATTATAACCCAATTGCTCAGAGGGCCTTGG
 TTCCCGTTACAGACCATGTAGGCTACATGATTCTGCAGAAAATATGTG
 25 AGGGTCATGCAGACAAAATGACAGAGTTGAATACCTCAAGAACATCAGTGT
 AGAGGAACACATTAGCCGGAACACAAGAAATCATCTTCAGATTCCATCTC
 AAATTAAAGGATTGGTGGACCAAGTAGAGAAGGGATCAGAGCAAATGTGGAA
 AACTTTCCGATTGATGTCATCACTTGTGAGTCTCAGGATCAGGCACAA
 GCTTGGACAGAAAGCCTCAAGATAACTGAGCAGATTGAAAGTCTAACAA
 30 GACAGCTCTCCCTGATCAGTTGGACTGATGATCCAGTTCCCTCTAGGAAGA
 GTTGGTTCCATGAATGCATCCACCTCTGCATCATCAAGTGTGATTCCC
 ATCAAGAGAGAAAACCTTACACAAGCACTAAAGCACTCGAACCCACC
 AACAAATTCCACATGGTAGCCTTGTGGATGGGTGGAGTAGGAAAGACT
 AGAATGATGCAAAGGCTGAAGAAGGCCCTGAAGAAAAGAAATTGTTAA
 35 TTATATTGTTAGGGCAGTTAGGGAAAAGACGGACCCCTTGCCATT
 AAGAAGCTATAGCAGATTACCTCGGTATACAACCTCAATGAAAAAAACTAAG
 CCAGCAAGAGCTGATAAGCTTCGTGAATGGTCAAAAAGAATTCAAGATGG
 AGGTAAGACTAAGTCTCATAGTACTTGACGATGTTGGCAATTAGTTG
 ATCTTGAAGATATTGGGTTAACGTCTTCCAAATCAAGGTGTCGACTTC
 40 AAGGTCTTGTGACATCACGAGACTCACAAGTTGCACTATGATGGGGGT
 TGAAGCTAATTCAATTAAACGTGGGCCTCTAACTGAAGCAGAAGCTC
 AAAGTCTGTTCCAGCAATTGTAGAAACTTCTGAGCCGAGCTCCAGAAG

ATAGGAGAGGATATCGTAAGGAAGTGTGCGGTCTACCTATTGCCATAAA
AACCATGGCATGTACTCTTAGAAATAAAGAAAGGATGCATGGAAGGATG
CACTTCGCGCATAGAGCACTATGACATTACAATGTTGCGCCCAAAGTC
TTTGAACAGAGCTACCACAATCTCCAAGAAGAGGAGACTAAATCCACTTT
5 TTTAATGTGTGGTTGTTCCGAAGACTTCGATATTCTACTGAGGAGT
TGATGAGGTATGGATGGGCTTGAAGCTATTGATAGAGTTATACGATT
AGAGAAGCAAGAACCAAGGCTAACACCTGCATTGAGCGACTGGTGCAGAC
AAATTGTTAATTGAAAGTGTGATGTTGGGTGTCAAGATGCATGATC
10 TGGTCCGTGCTTGTGTTGGTATGTTCTGAAGTCAGTCAGCATGCTTCT
ATTGTCAACCAGGTAATATGCCGAGTGGACTGAAAATGATATAACTGA
CTCTTGCAAAGAACATTCAATTAAACATGCAAGAGTATGCTAAGTTCCAG
GAGATTCAAGTTCCAAACCTAATGATTGAAACTTATGCATGGAGAT
AAGTCGCTAAGGTTCTCAAGACTTTATGAAGGAATGGAAAAGCTCCA
TGTATATCATACGATAAAATGAAGTACCCATTGCTCCTTGGCACCTC
15 GATGCTCCACCAACATTGGGTGCTCATCTCACTAAATGTTCAATTAAAG
ATGTTGATTGCTCTGTATTGAAATCTATCGAATCTGGAAGTGCTGAG
CTTGCTAATTCTCGCATTGAATGGTACCTTCCACAGTCAGAAATTAA
AGAAGCTAAGGTTACTGATCTGAGATTGATGGTCTCCGTATAGAA
CAGGGTGTCTGAAAAGTTAGTCAAACATTGAAGAATTATATTGGAAA
20 TGCATCTGGTTTATAGATGATAACTGCAATGAGATGGCAGAGCGTTCTG
ACAACCTTCTGCATTAGAATTGCGTTTTAATAACAAGGCTGAAGTG
AAAAATATGTCATTGAGAATCTGAACGATTCAAGATCTCAGTGGGACG
CTCTTTGATGGAAATATCAATATGAGTAGCCACTCATACGAAAACATGT
TGCAATTGGTACCAACAAAGGTGATGTATTAGACTCTAAACTTAATGGG
25 TTATTTGAAAACAAAGGTGCTTTTAAGTGTGCATGGCATGAATGA
TCTGAAGATGTTGAGGTGAAGTCGACACATCCTACTCAGTCCTCTCAT
TCTGCAATTAAAAGTTCTTATTATTCAAAGTGTAGAGTTGAGATAC
CTTTCAAACCTCAATCTGCAAACACTTGTCAAGACTTGAGCATCTAGA
AGTTGTGAATGCGAGAATATGGAAGAACTCATACATACTGGAATTGTTG
30 GAGAAGAGACAATTACTTCCCTAACGCTAACGTTTATCTTGAGTCAC
CTACCGAAGTTATCAAGTTGCCCCATAATGTCAACATAATTGGGCTACC
ACATCTCGTAGACTTGATACTAACGGCATTCCAGGTTCACAGTCATT
ATCCCGAGAACAGTTGCGAACATCTAGTTGTTGAAGGAAGAGGTAGAT
ATATGTTCTTATGTTAATACAATTAAATAATTTCACACAAATT
35 CATAATATATCTGTAATTGATTGTATGTTGTTATTGTTATATGTGG
CTATTAAGGGATGATTATTGCAAGGTTGTGATTCTAACGTTGGAGACAC
TTCAAATTGATGACATGGAGAACTTAGAAGAAATATGCCCTGTGAACCT
AGTGGAGGTGAGAAAGTTAAGTTGAGAGAGATTAAAGTGAGTAGCTGTGA
TAAGCTTGTGAATCTATTCCCGCGAACATCCCATTGCTCTGTTGCATCATC
40 TTGAAAGAGCTTAAAGTCAGAACATTGCGGTTCCATTGAATCGTTATTCAAC
ATTGACTTGGATTGTGCGGTGCAATTGGAGAAGAGACAACAAGAGCCT
CTTAAAGAACATCAACATGGAGAAATTAGGAAAGCTAACAGAGAGGTGTGGA
GGATAAAAGGTGCAGATAACTCTCATCACGGTTTCAAGCTGTT

GAAAGCATAAAGATTGAAAAATGTAAGAGGTTAGCAATATATTCACACC
TATCACCGCCAATTCTATCTGGTGGCACTTTGGAGATTCAGATAGAAG
GTTGCGGAGGAAATCACGAATCAGAAGAGCAGGTAACGCTTCAATTAA
CTTCTTAAGTAATTAAAGGACTAACCTCCTGTTTTGAATAATAAAGAG
5 GTGGGATGACTAAACCTGGGCATCACAATTGCAACAAAATGTTACAAACC
ATGAAACGTTCAAACCCATTCTGAATTAAAGGTTCAATACAAGTCATT
AAAAATATGGCTTAAATTCTTATATTATGTATCAACATGATTTC
TAGAGATCATTATTATAATAGTAAGTTAAAGCAATTAAATTAGAACT
AATTCTAACTTAGCTAATAATCGTTAAATGTAATAATTACTTTT
10 AGTGAAATAAGCAACGGATTAATAAGTTAACAACTTAAATGTCATTCC
TAACAAAAAAACTATTGGTCAGAAGAACCGTAATTCAAGATAACTAA
AATAAAAATATTGACATTCACTAACAGAGCATTCTTCTAAATATGAT
TGCAAATGAATAAAACTTAAATTATACAGAAAAGATTTTATATATGTT
ATACAAAATTACAAATTGAAACTGGATATGTTAATTACGGTTATAAT
15 TCTGGTATCACAAAGGGATATATAATAAAATTATTCTGTAGTCATT
TATAATTGTAAGTTATAACCCGTGGGAACCATGAGTTCTAAAATTAG
TTAAACTTCATAATAAAAATTATAATTATTATTATTAAATAATT
ATTAAATTAAGAGATGTATCAAAAATTAAAGTTATTATAACTTCAAATT
AACATATAATTAGAAAATATGATCATAACTTCCGCAACTCTCTT
20 GTATTAAAATGCCAGAGAAGCTCTAGTAYATTCTAAATCAAAGTC
CAAACATAATGAAGCATATAATTGTGAAAATCAATTAGCATTAGGTTT
TAAGAGTCACCAAATTCAAAGAGTAATCCAATGCTTCATTACCACTATG
GAGAAAATATTCTTAGTTAAATGAAATGAAAACAAACATTCAAAC
ATTGTTGCTTACTAAACCAAGACCCATTACTTAGCCAAGAGTTAACCA
25 AAAAAAATTACATTCATGTATCATTATTGACTAGATATATATGAACA
TGAAGGGAGTTTATAGAAAATATAATCATAGATATTCAACATAACTC
ATGGAATTCCCTAAAATAACCAAGTTATTCAAGAAATTACATCCAAGTC
ACCAAGAGAAAGTTAGCCTAGCATGGCTAAACTCAAGAAAATAAA
GGATTAGAAGTACCAACATGTAGTAAGAATCACAGTAAAAGATGATGTT
30 GTTCTTGATGTTCTCTAAGTTCTCAAGTCTCCAGTTGCTCCTAATAAT
GCAAGGGAGAGCCATTAAATTCTGATGTATTGATCCCTCAAAAGCTGCA
CCAACCTCCCTTAAATAACACTCAAAGCAAAATGACAAAATTGCCCTG
AAGGACCTATGCCGGTGCCTGCGGGTGGAGCTGAATACGAAAGGTC
TTGGTCTTGTGAGGGTGATGCTGTGCGGGTAGCTGTCATGCTTC
35 CGCGCGGTTCGCGCACATGTGCACAAGTGTGATGCATGGTGTACGTTCT
GAGTTTGAGCCTCGATGCTTAGTCCATTGGCCAATTGAGTCAAAT
CAGCTTATGACCCATTCTCAAGTTATCTCAAGTTATCTCAAGTT
AAGCCCAAATTGCCCTCTCAAATCATCCATAACTTCACAAAATCGCCCG
TTCATCTTAATCCGAATGCACAATTATTCTCCTGCTTCAAGCA
40 AGATACCACTTCTCATGCTTCATCCATCAATAGTACACTTCATGTATC
ATCTCTACTAGTTATTAGTCCACAATCCTTATTGCTCCAAATTAAAT
TATCTCATTAGTCCACTAGTTCCCTAAATTGCAATTAAAG
CTCACACAAATATTAAGTACCTGAAATGGTCATAAAATAACAAAAAGGAA

AATATGCATGAAGATTAACATAATGATGAACGAAATATGCTAAAATAGAC
TATAAAATGAAGTAAATAAAATGAAATTATCGCACTCCGACCACCCCTAT
AGGCTTGTAGTCATCCACCCCTCATTCTGTACCAATATGGGATGGAA
ACATCATTAATTAAGCCAAAAACTAACATATAAGGGGTGAGTGACAAAG
5 GTAAGTACTAAAGATGAAAATAATCCATTTTYTGTATATACACAACAC
ACACATAGGGGCAGACGTAGGATTTCATAGTACAGATTGTTGGCACA
TAAGTGTGCTGGTACACTTTTTCTTACGTAGTGGCACAACAG
TAGAAAAAACGARAAATTCAAATTTTACAATGTGTSTAAAAAAAAYA
GTGGTTGTTGGTGCCACTATGGACACCAAAGTTGAAC TGCCCCTGCGCGC
10 RCACACACACACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
ARAGWAWGRRRGAKAKARMCSMSYTTGGGATGTGATACTTCTTTAGGAA
AATGGAGTTATATCTTGATATTGTATTTTTAATGTAATTATATATT
TAATCATTAGTTATAAGTTTATTATTTGATATGAAAAAAAAGT
CTTTATACATTGGATTTAACATAAAAATCCAACAATATTAATCAAAAAG
15 ACCAMACATGTGGACAMWTATGTATATAAWTAATTACAATAGTCTTAG
GAATAGNATTATATATAATTAAATTCTCAATGGCTTAGGAATAGTAAG
TTCTTATATTCAAACCTTNGCCACAATTCTTGKTTACTWGACACTY
CCTCTCTCTAATTATATATATATATATATATATATATACACA
CACACACACACACTAGATGTGTGCCCGCGAAAGCAGTGACGTNNNG
20 AGAANACTTCTTAAGCATAAATAATTATTATTTTATTGGGTATTA
TATAATAAAAAATTACAACCTTAAATAAAATATTATGTTTACTTTA
TATTATATTGCTGTACTATTAAATATAATAAAATTAAATTATGTCT
AATTATGAAATGTAATTAAATTAAACATGAATTAAATATTAA
ATTTCAAGTTGCTCAAATTGAGTTCTAATTATTTTTAATT
25 GTATTCAAACCTTGGTAAGTATTAAAGAATTATTTATGCACAATTGATT
TATACAAAAAACTTGTAACTTACATCTTAAATTCAAGATATAACTA
ACATGTTTACAATATATATATATANATATATATATATATATAT
ATATATATATATAGTAAAGCGCANAGGTACAGGNANAGANTATT
TCTATTATTCTACGTTTGCCACAAAGTTGAACACTTGCCACTTTT
30 GTCCTCCTAACCTTCAATGTTTGCAGACAAAAGTCCAAAACCTTG
CCACTTGTACATTCTCAACTTTCACCGCATTAGTTGTGGAGTTGGC
AGTTTGGTCCCCCTAACCTCGATATTCTCTGCTAGCCAAAAGGGT
TCCAGAGTTCACANTTTGGTCCCTGACAATAACCAATGTGAGATGTC
AAAATTTGCCACATTAGTTGTGGAGTTGTCCCTTGGTCCCCCACA
35 TTCGATATTCTACTACGACCTTATTCTCAAATAACAACACGTATA
TTTAAATTACCAATGATAGAAATAGATATCAAATAAAAGTATTGTAACACC
GTGTAAGAACGGTGCTACTATAGTAAAATAACATTCAAAGTACGAT
GTCCTAATTGGAAAAAGAGTTAAAAAATAACAACACTAGGGCGAGTT
TTTTACAAGTTGTATCAAATCATATAAAAGTGGAAACGGTGA
40 CCACATTAACCAGAAATGTAATTATTCTTGATTTGATAATTAAAT
ATTGTGTTGATCTATGTATTAAAGTAAACAACAAAGAACATAATCC
AAAACCTAAATTGCAAGTCTGCCAATTCTCTATCACTAGTCGTAC
TTACGATGGCGTTACGTCGCTCTCACTTCTACAACCCTTGTGCTA

CTCATTACAATAACGAAAAGTTGAATATCCATATATTATTTGGATGTGG
AATTGAACAAATCTCGTCAAATTGGATTTGATGGATTTGAGTAG
AAGTTGGCAGAACGGGAATGATGGCTGCAAGTGGTATAAAGTGTAT
TCTGAGTTATTACTATATATGTAGCCTCTTACAACGACCAAGGTTCTT
5 CCAGGTACCATTGATCTTTAGAACCCAGTGTCTGAAACACCCCTGAT
TTGGATCAAATATACCAACAACCTTAAAAACTGATTAATCAATTGTT
TTCTTCATCTTGATAACAAGTGGATGATTTCTACTTAGATTAACCTGA
AAAAAAAGGTCCATGTGCCTGGATCTGGTAAATGAAGATGGAAGG
GAGAGCTGACTTTAAAGACACAAACACGTACCCATATCTTTATTTATT
10 TTAAATTGCTTTTCCTATTCTTCTTGTATCTCCAGATGGTAT
GTGGTGTGGATAATTACACATAGAGATTGGAACGACTGTGTTAGAG
AGGACGTGGCTGGGTTGAGGATGGTTATGGCTGGCGAGTTCATTT
ATATAAACAAACAAATATATAAAACAAGGGTAAAATGCCATCTTATAT
GTATTAAACCGTCCTTTTATTTTTATTAAATTAAAGAAGG
15 GGTATACCAGTGTCAAGCCTTATTCCAACCAGGCAACCAAGTCAAATAG
GGACTTAGGTTGGAAACAGTCCGTGAGACCGTGACTTGGATGGTA
GATAAATTAGTAAACTAACCTTCATTAAACCTACCTTTCTTATTAA
ACTCAATTCAACCTAAATTCTGATTCTGTTGAAAATAAGTTGCATCT
TTATGTTGTATTATCCTGTTGCATAGGATCCTAGCATCTTTAATAAT
20 TTAATTGAAGGTGAAAGATCCAACATTTTTAGCTGTGGATTTC
TCATTGCAACTGTTCTGAAAAAAACCTAAATCAAATAACCA
TTTCAAATCCAAAATTATAAGAGAGAATTGTAATGGACGTGAAATCGT
AAATCATTAAACACAGTCAGTACACAAGTTGCTAATTACATTCTGCTG
TGCAGATTGAAATTCTATCAGAGAAAGAGACATTACAAGAAGTCACTGAT
25 ACTAAATATTCTAATGATGTTGATTATTCCATCCTGCTCATGCACTC
TTTCATAACCTCCATAAAACTTAAATTGGAGAGAGTTAAAGGAGTGGAGG
TGGTGTGAGATAGAGAGTGGAGAGTCCAACAAGTAGAGAATTGGTAACA
ACTCACCATAACCAACAACATCCTATTATAACTTCCACCTCCAGGAATT
GGATCTAAAGTTTATGGACAACATGAGTCATGTGGAAGTGCAGCAACT
30 GGAATAAAATTCTCACTCTCCAAAACAACAAATCAGAATCCCCATTCCAC
AACCTCACAAACCATACACATGTTAGCTGAGCAGATTAAGTACTGTT
TTCGCCTCTCATGGCAGAACCTTCCAAACCTAAAGGATATCTGGATAA
GTGGGTGTAATGGTATTAAAGAAGTTGTTCAAAGAGAGATGATGAGGAT
GAAGAAATGACTACATTACATCTACCCACACAACCACCATTTGTTCCC
35 TCATCTTGATTCTCACTCTAAGACTACTGGAGAATCTGAAGTGTATTG
GTGGAGGTGGTCCAAGGATGAGGGAGCAATGAAATATCTTCAATAAT
ACCACGTCAACTACTGCTGTTCTGATCAATTGAGGTATGCTTGTACA
TATTCAATTATTAAATTCTTTCTTGTCAATATTCTATAAAT
AATACATTTATACCCACTACTAAGATAATAATTACCTAGAGGGATGG
40 ATGCTATGACACAGCTGCTACACTTCAGAAACTCTAGTAAGGGCAGTTAT
GGAAGTTCAATAAAATGATAATGGCATCTTGATGGTAATATAGGCAA
TTTAAAGTTTATTTCTGTTAAAGCAGTATTAGCAAGTACTGGCCAGTAG
GAGAGGAGAATATCACCTTGTGAAAATCTGGTCATTGTACCCAAGAAT

TTAGTTAACATTTAGATATCAGGGGACATCAGGTGACAGATAT
TGTAGAATAGAACAAATATAATATTACCCAAAATTTCTAAGGT
TATTCTGTTAAATATGTGCTTCTGATTCAATTGCATTCTAT
ATTTAGGTGGTAAAGTGATTGCTCTCAATAAATCCGAAATTAA
5 AAAAAAAACAAAAGTAAATTTGATATGGAGAGCACTGGTATCA
TTTAGTATATAAAAAACTAGATTGAATTAAGTTCTATATAAAAGC
TGTGTATATAGTTAATTAGTTACATCATTTCCATGTGGTGTGCA
GTTGTCTGAAGCAGGTGGTGGAGTTATGCCAATACGCTAGAG
AGATAGAGATATCTAAGTGTAAATGTATTGTCAAGTGTGATTCCATGTTAT
10 GCAGCAGGACAAATGCAAAAGCTCAAGTGCTGAGAGTAACGGGTTGTGA
TGGCATGAAGGAGGTATTGAAACTCAATTAGGGACGAGCAGCAACAAAA
ACAGAAAGGGTGGTGGTGTGAAGGAAATGGTGAATTCCAAGAGTAAAT
AACAAATGTTATTATGCTTCCAATCTAAAGACATTGAAAATCTACATGTG
CGGGGGTTGGAACATATATTCACATTCTCTGCACITGAAAGCCTGACAC
15 AGCTCCAAGAGTTAAAGATAGTGGTTGCTACCGAATGAAAGTGTGATTGTG
AAGAAGGAAGAAGATGAATATGGAGAGCAGCAAACAACAACAAC
AACGAAGGGGGCATCTTCTTCTTCTTCTCTAAGAAGGTTG
TGGTCTTCCCCGCTAAAGTCATTGAACATTCAATCTACAGAGCTG
GTAGGATTCTCTGGGATGAATGAGTTCCGGTTGCCTCATTGGAAGA
20 AGTTACCATCAAGTATTGCTAAAAATGATGGTGTGAGCTGGTGGGT
CCACAGCTCCCCAACTCAAGTATATACACACAAGATTAGGAAACACT
CTTGATCAAGAATCTGGCCTTAACCTTCATCAGGTATATATATTCCTT
TAATTGGCATGATCTAATTAAGAAAGATATCATTCTGCCAAGTAAATT
ACTTCAAACACATTACACTGGTTCACTCTAAGTTATGTTGTTCTAGG
25 AAGGCCAAATGGAAAGCAAGATAGGGAAAAATAGTGTATTCAGTGG
AAGGGTATTTAGGTATTTCTGTCAAAAGTTGTTATTGCAAGGCTTTA
GTACCTGGAATCGTGTGGAGGAGCGTTATTCTGATTGCTTGT
TCTTATCATTCTTAGCCTCTCGAACAGCTAGAAACCCCTTAATC
TTTGATTAAATGACAAATTTCCTGTTACTCTATTGATTGTTG
30 TTCTTCATGGTCTAAGTGAGTTATTGGCTCATCTGTTACTCTTTGAT
TGTATTTCATATCATGTTGCTTGAATCAAGCTTCCATTCAA
CCAGGGAAAAGGTCAAAAGTAACCTACTTATGAGATCAAAACAGCAA
CCCATCGGATAACTTTAGTGGAGTTAATAGTACAATTACATTGTGA
TTAATAATTATAATCTGTATTAAATTCAATTAAAATTGGTACAGCACAT
35 ATATGACATTAAAGGTTGTTTGACATATATATGCCCTGGC
GTTTCTTATTGGACATGCAGACCTCATCCAAAGTTATACGGTGACA
CCTCGGGCCCTGCTACTTCAGAAGGGACAACCTGGTCTTCATAACTTG
ATCGAATTAGATATGGAATTAAATTATGATGTTAAAAGATTATTCCATC
CAGTGAGTTGCTGCAACTGCAAAAGCTGGAAAAGATTGAGTAGTT
40 GTTATTGGGTAGAGGGAGGTATTGAAACTGCATTGGAAGCAGCAGGGAGA
AATGGAAATAGTGGATTGGTTGATGAATCGTCACAAACTACTAC
TACTACTCTTCAATCTCGAAACCTCAGAGAAATGAAGTTGCATTTC
TACGTGGTCTGAGGTATATGGAAGAGCAATCAGTGGACAGCATTGAG

TTTCCAAACCTAACAAAGAGTCATATAAGTAGGTGAGAAGGTTAGAACAA
 TGTATTTACTAGTCCATGGTGGTAGTCTATTGCAACTCCAAGAGCTAG
 ATATTAGTTGGTGCACCATAATGGAGGAGGTGATTGTTAAGGATGCAGAT
 GTTTCTGTGAAGAACAAAGAGAGAGAACATCTGATGGCAAGACGAATAA
 5 GGAGATACTGTGTTACCTCGTCTAAAATCCTGAAATTAAAATGCCCTC
 CATGTCTTAAGGGGTTAGCTTGGGAAGGAGGATTTCATCCATTAA
 TTGGATACTTAGAAATCTACAAATGCCAGCAATAACGACCTTCACCAA
 GGGAAATTCTGCTACTCCACAGCTAAAAGAAATAGAAACAAGATTGGCT
 CGTTTATGCAGGGAAAGACATCAACTCCTCTATTATAAAAAGATCAAAC
 10 AACAGGTAAATCAGATCTTGTGCTTAATAATTCTTAAACTACATTG
 AAAAGCTTCATGCAAGTTTTGTATATTGTCAAAAACCGCAACCTA
 CATTTCAGCTTATATTATGTACTTATGCAGGAGTCAAACAAACT
 CTGATTAATGTGAAGTGAATATTAAAGGTAAATTATATTTCATGTTCT
 AGTTGCCTATTAAATTAAATGGCCTTTAGTCRTGATTGGATGTAGTY
 15 WTCATGATGATGTGAATCTTCTAATACCCCATTCAATTGTTGGATGAATG
 TTGACTCTATGTCAGGATGAATATTCAAGGGAAGAATTGTTCATCATATG
 AAGGACATTAAAGAACATGGATGCTATGAAGATGTTGGAARAC

RG2S deduced polypeptide sequence (SEQ ID NO:125)

20 MSDPTGIAGAIINPIAQRALVPVTDHVGYMISCRKYVRVMQTKMTELNTSRISVEEH
 ISRNTRNHLQIPSQIKDWLDQVEGIRANVENFPIDVITCCSLRIRHKLGQKAFKITEQI
 ESLTRQLSLISWTDDPVPLGRVGSMNASTSASSDDFPSREKFTQALKALEPNQQF
 HMVALCGMGGVGKTRMMQRLKKAAEKKLFNYIVRAVIGEKTDFAIQEAIADYL
 GIQLNEKTKPARADKLREWFKNSDGGKTKFLIVLDDVWQLVDLEDIGLSPFPNQG
 25 VDFKVLLTSRDSQVCTMMGVEANSIINVGLTEAEAQSLFQQFVETSEPELQKIGED
 IVRKCCGLPIAKTMACTRNRKDAWKDALSRIEHYDIHNVAPKFETSYHNLQE
 EETKSTFLMCGLFPEDFDIPTEELMRYGWGLKLFDRVYTIREARTRLNTCIERLVQT
 NLLIESDDVGCVKMHDLVRAFVLGMFSEVEHASIVHGNMPWEWTENDITDSCKRIS
 LTCKSMSKFPGDFKFPNLMILKLMHGDKSLRFPQDFYEGMEKLHVISYDKMKYPLL
 30 PLAPRCSTNIRVLHLTKCSLKMFDSCSIGNLSNLEVLSFANSRIEWLPSTVRNLKKLR
 LLDLRFCDGLRIEQGVLSLVKLEFYIGNASGFIDDNCNEMAERSDNLSALEFAFF
 NNKAEVKNMSFENLERFKISVGRSFBDGNINMSSHSYENMLQLVTNKGDVLDLSDKLN
 GLFLKTKVLFLSVHGMNDLEDVEVKSTHPTQSSFCNLKVLIIISKCVELYLFKLNL
 ANTLSRLEHLEVCECENMEELIHTGICGEETITFPKLKFLSQLPKLSSLCHNVNIIG
 35 LPHLVDLILKGIPGFTVIYPQNKLRTSSLKEEVVIPKLETQLQIDDMENLEEIWPCELS
 GGEKVKLREIKVSSCDKLVNLFPRNPMSSLHHLEELVKVNCGSIESLFNIDLDCVGA
 IGEEDNKSLLRSINMENLGKLREVWRIKGADNSHLINGFQAIVESIKIEKCKRFSNIFT
 PITANFYLVALLEIQEGCGGNHESEEQIEILSEKETLQEVTDTNISNDVVLFPSCLMH
 SFHNLHKLKLERVKGVENVFEIESESPTSRELVTTHHNQHQPIILPNLQELDLSFMD
 40 NMSHVWKCSNWNKFFTLPKQQSESPFHNLTTIHMFSCRISIKYLFSPMAELLSNLK
 DIWISGCNGIKEVVSKRDDEDEEMTTFTSTHTTILFPHLDSTLRLLENLKCIIGGGG
 AKDEGSNEISFNNTTATTAVLDQFELSEAGGVWSLCQYAREIEISKCNVLSSVIPCY

AAGQMQLQVLRVTGCDGMKEVFETQLGTSSNKNRKGGDEGGNGIPRVNNV
MLPNLKTLCIYMCGLHEHIFTSALESLTQLQELKIVGCYGMKVIVKKEEDEYGEQ
QTTTTTCKGASSSSSSSSKKVVVFPRLKSIELFNLPELVGFFLGMNEFRLPSEEV
5 IKYCSKMMVFAAGGSTAPQLKYIHLRLGKHTLDQESGLNFHQTSFQSLYGDTSGPA
TSEGTTWSFHNLIELDMELNYDVKKIIPSSELLQLQLEKIHVSSCYWVEEVFETAL
EAAGRNGNSIGFDESSQTTTTLFNLRLNREMKLHFLRGLRYIWKSQNWTAFEF
PNLTRVHISRCRRLEHVFTSSMVGSLQLQELDISWCNHMEEVIVKDADSVVEEDK
ERE SDGKTNKEILVLPLKSLKLKCLPLKGFSLGKEDFSFPLLDTLEIYKCPAITTFT
KGNSATPQLKEIETRFGSFYAGEDINSSIJKRSNNRSSNKTINVK.ILK

10

RG2T polynucleotide sequence (SEQ ID NO:126)

GGAAGACGACAATGGTGCAACGGTTGAAGAAGGTTGTGAAAGATAAGAAG
ATGTTCCATTATATTGTCGAGGTGGTTGTAGGGGCAAACACTGACCCCAT
TGCTATCCAGGATACTGTTGCAGATTACCTCAGCATAGAACTGAAAGGAA
15 ATACGAGAGATGCAAGGGCTTATAAGCTTCGTGAATGCTTAAGGCCCTC
TCTGGTGGAGGTAAGATGAAGTTCTAGTAATTCTTGACGATGTATGGAG
CCCTGTTGATCTGGATGATATCGGTTAACGTTCTTGACATCACGAAACAGTGA
TTGACTTCAAGGTCTGCTGACATCACGAAACAGTGAATCTGCATGATG
ATGGGAGCTAGTTAATTTCACCTCAATATGTTAACAGACGAGGAAGC
20 ACATAATTTCGTCGATACGCAGAAATTCTTATGATGCTGATCCCG
AGCTTATTAAGATAGGAGAACGATTGAGAGAAATGTGGTGGTTACCC
ATTGCCATAAAACATGGCCGTTACTCTTAGAAATAACGCAAAGATGC
ATGGAAAGATGCACTTCTCGTTAGAGCACCGTGACACTCATAATGTTG
TGGCTGATGTTCTAAATTGAGCTACAGCAATATCCAAGACGAGGAGACT
25 CGGTCGATTTCGCTATGTTGATGCTGTTGTTCTGAAGACTTGTATATTCC
TACCGAAGACTTAGTGAGGTATGGATGGGATTGAAAATATTACCAAGAG
TGTATACTATGAGACATGCAAGAAAAAGGTTGGACACGTGCATTGAGCGG
CTTATGCATGCCAACATGTTGATAAAAAGTGATAATGTTGGATTGTCAA
GATGCATGATCTGGTTCTGCTTGTGTTGGCATGTTATCTGAAGTCG
30 AGCATGCATCAATTGTCACCATGGGGATATGCCAGGGTGGTTGAAACT
GCAAATGATAAGAACAGCTTGTGCAAAAGAATTCTATTAACATGCAAAGG
TATGTCTGCGATTCTGAAGACCTCACGTTCAAACCTCTCGATCCTGA
AATTAATGGATGGAGACGAGTCACTGAGGTTCTGAAGGCTTGTGAA
GAAATGGAAAACCTCAGGTTATATCATATGATAACATGAAGCAGCCATT
35 TCTTCCACAAATCACTCAATGCTCCAATGTCGAGTGCTCATCTCCATC
ACTGCTCATTAATGTTGATTGCTCTTCTATTGAAATCTTTGAATCTC
GAGGTGCTCAGCATTGCTAATTCTGCCATTAAATTGTTACCCCTCCACTAT
TGGAGATCTGAAGAACGTAAGGCTCCTGGATTGACAAATTGTGTTGGTC
TCTGTATAGCTAATGGCGTCTTAGAAATTGGTCAAACCTGAAGAGCTT
40 TATATGAGAGTTGATGATCGAGATTGTTGTGAAAGCTGATGACAG
CAAGACCATTACCT

RG2T deduced polypeptide sequence (SEQ ID NO:127)

KTTMVQRLKKVVDKKMFHYIVEVVVGANTDPIAIQDTVADYLSIELKGNTNDAR
 AYKLRECFKALSGGGKMKFLVILDDVVSPVDLDDIGLSSLPNQGVDFKVLTSRNS
 DICMMMGASLIFNLNMLTDEEAHNFFRRYAEISYADPELIKIGEAIVEKGGLPIA
 5 KTMAVTLRNKRKDAWKDALSRLEHRDTHNVVADVLKLSYSNIQDEETRSIFLLCG
 LFPEDFDIPTEDLVRYGWGLKIFTRVYTMRHARKRLDTIERLMHANMLIKSDNVG
 FVKMHDLVRAFVLGMLSEVEHASIVNHGDMPGWFETANDKNSLCKRISLTCKGMS
 AIPEDLTFPNLSILKLMGDDESRLFPEGFYGEMENLQVISYDNMKQPFLPQLQCSN
 VRVLHLHHCSLMFDCSSIGNLLNEVLSIANSAIKLLPSTIGDLKKLRLLDLTNCVGL
 10 CIANGVFRNLVKLEELYMRVDDRDSFFVKADDSTKTIT

RG2U polynucleotide sequence (SEQ ID NO:128)

GCCTTGTGTGGATGGGTGGAGTGGAAAGACCACTGTGATGAAGAACGCT
 GAAGGAGGTTGTGGTAGGAAAGAAACTGTTTAATCATTATGTTGAGGCAG
 15 TTATAGGGAAAAGACAGACCCCCATTGCTATTCAACAAGCTGTTGCCAG
 TACCTTGGTATAAGCTAACCGAAACCACTAAACCAGCAAGAACTGATAA
 GCTCCGTACATGGTTGCAAACAACACTCAAATGGAGGAAAGAAGAAGTTCC
 TGGAATAACTAGACGATGTATGGCAACCAGTTGATTGGAAGATATTGGT
 TTAAGTCGTTTCCAATCAAGATGTTGACTTCAAGGTCTGATTACATC
 20 ACGGGACCAATCAGTTGCACTGAGATGGGAGTTAACAGCTGATTAGTT
 TCAAGGTGAGTGTCCCTGGAGGAAGCGGAAGCACACAGTTGTTCCCTCAA
 TTTTAGAACCTCTGATGATGTCGATCCTGAGCTAACAAAATCGGAGA
 AGAAATTGTAAGAAGTGTGAGACTACCCATTGCTATCAAACCATGG
 CCTGAACCTCTAGAAGTAAAAGTAAGGATACATGGAAGAATGCCCTTCT
 25 CGTTTACAACACCATGACATTAACACAATTGCGTCTACTGTTTCCAAAC
 TAGCTATGACAATCTCGAAGACGAGGTGACTAAAGCTACTTTTGCTTT
 GTGGTTTATTCCGGAGGACTTCATATTCCCTACCGAGGACCTATTGAGG
 TATGGATGGGATTGAAGTTATTCAAGGAAGTAGATACTATACGAGAAC
 AAGATCCAAGTTGAAAGCCTGCATTGAGCGGCTCATGCATACCAATTGT
 30 TGATCGAAGGTGATGATGTTAGGTACGTTAAGATGCATGATCTGGTGCCT
 GCTTTGTTGGATATGTTCTAAAGCCAGCATGCATCTATTGTCAA
 CCATGGTAGTAGTAAGCCAAGGTGGCCTGAAACTGAAAGTGATGTGAGCT
 CCTCTGCAAAAGAATTCTTACATGCAAGGGTNTG

35 RG2U deduced polypeptide sequence (SEQ ID NO:129)

ALCGMGGVGKTTVMKKLKEVVVGKKLFNHYVEAVIGEKTDPPIAIQQAVA
 EYLGISLTETTKPARTDKLRTWFANNNSNGGKKFLVILDDVVWQPVDLEDIGLSRFPNQDV
 FKVLITSRDQSVCTEMGVKADLVLKVSVLEEAEAHSLFLQFLEPSDDVDPELNKIGE
 EIVKKCCRLPIAKTMA.TLRSKSKDWTKNALSRLQHHDINTIASTVFQTSYDNLEDE
 40 VTKATFLLCGLFPEDFNIPTEDLLRYGWGLKLFKEVDTIREARSKLACIERLMHTN

LLIEGDDVRYVKMHDLVRAFVLDMFSKAEHASIVNHGSSKPRWPETESDVSSCKR.
ISLTCKG?

RG2V polynucleotide sequence (SEQ ID NO:130)

5 CTGTGGAAGACACGAATGATSAAGAAGCTGAAGGAGGTCGTGGAACAAAA
 GAAAATGTTCAATATTATTGTTCAAGTGGTCATAGGAGAGAAGACAAACC
 CTATTGCTATTCAAGCTGTAGCAGATTACCTCTATTGAGCTGAAA
 GAAAACACTAAAGAAGCAAGAGCTGATAAGCTCGTNAATGGTTGAGGA
 CGATGGAGGAAAGAATAAGTTCTTGTAAATTGATGATGTATGGCAGT
10 TTGTCGATCTGAAGATATTGGTTAACGCCTCTGCCAAATAAAGGTGTC
 AACTTCAAGGTCTTGTGACGTTAACAGAGATTACATGTTGCACTCTGAT
 GGGAGCTGAAGCCAATTCAATTCTCAATATAAAAGTTTAAAAGATGTTN
 AAGGACAAAGTTGTCGCCAGTTGCTAAAAATGCAGGTGATGATGAC
 CTGGATCCTGCTTCAATGGGATAGCAGATAGTATTGCAAGTAGATGTCA
15 AGGTTGCCCATGCCATCAAAACCATGCCTTAAGTCTAAAGGTAGAA
 GCAAGCCTGCGTGGGACCATGCGCTTCTCGTTGGAGAACCATAGATT
 GGTAGTGAAGAAGTTGTCGTGAAGTTTAAAATTAGCTATGACAATCT
 CCAAGATGAGGTTACTAAATCTATTTTWTACTTGCTTATTCTG
 AAGATTTGATATTCCATTGAGGAGTTGGTGGAGGTATGGGTGGGCTTG
20 AAATTATTATAGAACAAAAACTATAAGAGAACAGCAAGAACAGGCTCAA
 CACCTGCACTGAGCGGCTTAGGGAGACAAATTGTTATTGGAAGTGATG
 ACATTGGATGCGTCAAGATGCACGATGTGGTGCCTGATTTGTTGGTAT
 ATATTCTCAGAAGTCCAGCACGCTCAATTGCAACCAGGTAAATGTGTC
 AGAGTGGCTAGAGGAAAATCATAGCATCTACTCTTGTAAAAGAATTCT
25 TAACATGCAAGGGTATGTCGAGTTCCAAAGACCTCAAATTCCAAAC
 CTTCATTTGAAACTATGCATGGAGATAAGTCNTGAGCTTCCTGA
 AGACTTTATGAAAGATGAAAAGGTCAGGTAATATCATATGATAAAT
 TGATGTATCCATTGCTTCCCTCATCACTGAAATGCTCCACTAACGTTCGA
 GTGCTTCATCTCATTATTGTCATTAAGGATGTTGATTGCTCTCAAT
30 TGGTAATCTCTAACATGGAAGTGCTCAGCTTGCTAATTCTAACATTG
 AATGGTTACCATCTACAATTGAAATTGAAAGAAGCTAAGGCTACTAGAT
 TTGACAAATTGAAAGGTCTCGTATAGATAATGGGTCTTAAAAAATT
 GGTCAAACCTGAAGAGCTTATATGGGTGTTAATGTCGTATGGACCGAGG
 CCGT
35

RG2V deduced polypeptide sequence (SEQ ID NO:131)

LWKTRM?KKLKEVVEQKKMFNIIVQVVIGEKTNPPIAIQQAVADYLSIELKENTKEAR
ADKLR?WFEDDGKKNKFLVILDDVVWQFVDLEDIGLSPLPNKGVNFKVLLLRDSH
VCTLMGAEANSILNIKVLKD?GQSLFRQFAKNAGDDDDPAFNGIADSIASRCQGL
40 PLAIKTIALSLKGRSKPAWDHALSRLENHKIGSEEVREVFKISYDNLQDEVTKSIF?L
 CALFPEDFDIPIEELVRYGWGLKLIEAKTIAREARNRLNTCTERLRETNLLFGSDDIG

CVKMHDVVRDFVWYIFSEVQHASIVNHGNVSEWLEENHSIYSCKRISLTCKGMSEF
 PKDLKFPNLSILKLMHGDKS?SFPEDFYGKMEKVQVISYDKLMPILLPSSLECSTNV
 RVLHLHYCSLRMFDCSSIGNLLNMEVLSFANSIEWLPSTIGNLKKLRLLDLTNCKG
 LRIDNGVLKNLVKLEELYMGVNVRMDQAV

5

RG2W polynucleotide sequence (SEQ ID NO:132)

TTGGGAAAGAGACAATGATGAAGAATTGAAAGAGGTTGTGGTTGAAAAGA
 AAATGTTAACATTATGTGGAGCGGTTATAGGGGAGAAGACGGACCCC
 ATTGCTATTCAAGCAAGCCGTTGCAGAGTACCTGGTATAATTCTAACAGA
 AACCACTAACGGCAGCAAGAACCGATAAGCTACGTGCATGGCTTCTGACA
 ATTCAGATGGAGGAAGAAAGAAGTTCTAGTAATACTAGACGATGTATGG
 CATCCGGTTGATATGGAAGATATTGGTTAACGATGTTCCCAAATCAAGG
 TGTCGACTTCAAGGTCTGATTACATCACGGGACCAAGCTGTTGCACTG
 AGATGGGAGTTAACAGCTGATTCAAGGTGAGTGTCCTAGAGGAA
 GCTGAAGCACAAAGCTTATTCTGCCAACCTTGGGAAACCTCTGATGATGT
 CGATCCTGAGCTCCATCAGATTGGAGAAGAAATTGTAAGGAAGTGTG
 GTTACCCATTGCAATAAAACCATTGGCCTGCACTCTTAGAAGTAAAGC
 AAGGATACATGGAAGAATGCACTTCTCGTTACAACACCATGACATTAA
 CACAGTCGCGCCTACTGTTTCAAACACAGCTATGACAATCTCCAAGATG
 20 AGGTGACTGGAGATACTTTTGCTATGTGGTTGTTCCGGAGGACTTC
 GATATTCCACTGAAGACTTATTGAAGTATGGATGGGGCTTAAATTATT
 CAAGGGAGTGGATTCTGTAAGAGAAGCAAGATAACCAAGCTGAAACGCGCTGCA
 TTGAGCGGCTCGTGCATACCAATTGTTGATTGAAAGTGATGTTGGG
 TCGTCAAGTTGCACGATCTGGTGCCTTATTGTTGATGTTGG
 25 TAAAGCGGAGCATGCTTCGATTGTCAACCATTGGTAGTAGTAAGCCTGGGT
 GCCCTGAAACTGAAAATGATGTGATCAGGACCTCCTGAAAAGAATCTCA
 TTAACATGCAAGGGTATGATTGAGTTCTAGTGACCTCAAGTTCCAAA
 TGTCTTGATTTAAAACCTATGCATGGAGATAAGTCGCTAAGGTTT

30

RG2W deduced polypeptide sequence (SEQ ID NO:133)

WERDNDEELKEVVVEKKMFNHYVEAVIGEKTDPIAIQQAVAELYGIILTETTKAAR
 TDKLRAWLSDNSDGGRKFLVILDDVWHPVDMEDIGLSRFPNQGVDFKVLITSRD
 QAVCTEMGVKADSVIKVSLEEAEASQLFCQLWEPSDDVDPELHQIGEEIVRKCCG
 LPIAIKTMACTLRSKSKDWTKNALSRQLQHHDINTVAPTVFQTSYDNLQDEVGDTF
 35 LLCGLFPEDFDIPTEDLLKYWGWLKLFKGVDSVREARYQLNACIERLVHTNLLIESD
 VVGCVKLHDLVRAFILDMFCKAEHASIVNHGSSKPGWPETENDVIRTSCKRISLTCK
 GMIEFSSDLKFPNVLILKLMHGDKSLRF

RG5 polynucleotide sequence (SEQ ID NO:134)

40 GGGGGGGTGGGAAGNCGACTCTAGCCCAGAAGNTCTATAATGACCATAA
 AATAAAAGGAAGCTTAGTAAACAAGCATGGATCTGTGTTCTCAACAAT

ATTCTGATATTCAGTTGAAAGAACGCCCTCGAACATCGGTGGAT
TATAAGCATGATGAAACTGTTGGAGAACCTAGCAGAAGGCTGCAATAGC
TGTGAAAATGCAAGTTCTTCTTGTGGATATTGGCAACATG
AGGTGTGGACTAATTACTCAGAGCCCCATTAAACACTGCAGCTACAGGA
5 ATAATTCTAGTAACAACCGTAATGATACAGTTGCACGAGCAATTGGGGT
GGAAGATATTGATCGAGTAGAATTGATGTCAGATGAAGTAGGATGGAAAT
TGCTTTGAAGAGTATGAACATTAGCAAAGAAAGTGAAGTAGAAAACCTA
CGAGTTTAGGGGTTGACATTGTCGTTGTGGTGGCTCCCCCTAGC
CTT

10

RG5 deduced polypeptide sequence (SEQ ID NO:135)

GGVGKTLAQK?YNDHKIKGSFSKQAWICVSQQYSDISVLKEVLRNIGVDYKHDET
VGELSRRLAIAVENASFFLVLDIWIWQHEVWTNLRAPLNTAATGIIIVTTRNDTVA
RAIGVEDIHRVELMSDEVGWKLKSMNISKESEVENLRVLGVDIVRLCGGLPLAL

15

RG7 polynucleotide sequence (SEQ ID NO:136)

GGTGGGGTGGGAAGACAAACGGGCACAAGGAGGCGACTGCCAATACCC
GACTTTATTGATAGAGATGACGAGTCTTATTTCCTACTACTATAGGGA
GGATATTGGTTGCGCGAGACGATTGCGCGAAGGGATTCTATCCTT
20 CTTTTTCCCGAAGACTTCGTTCCGGAGGACGGGCTATATTCCCTTA
ATATTAGTCTAGCCCAGTCTAGGCCAACCATATGGCGATGCGGTAGACCT
CCCAGAGATAGATACTTGATCTTAGAGGATTCACACGTTCAATGGTGGAA
ACTTAAGGAACCGGCTAAGAGTGACTAAACGGAAAAACCCCTATTCC
ATAGCCTCATCCGGTCGAGGCATTAACAATCCATCCAATCCTCTTCC
25 TTTGGTCTACTCTAATGATGTGCCGTTGTTGGGAATATCTCTTAT
ACCGACGATTATATGGGGATTGCCACTAGCGTTG

The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid construct comprising an RG polynucleotide which encodes an RG polypeptide having at least 60% sequence identity to an RG polypeptide from an RG family selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.
5
2. The nucleic acid construct of claim 1, wherein the RG polynucleotide encodes an RG polypeptide comprising an leucine rich region (LRR).
10
3. The nucleic acid construct of claim 1, wherein the RG polynucleotide encodes an RG polypeptide comprising a nucleotide binding site (NBS).
15
4. The nucleic acid construct of claim 1, wherein the polynucleotide is a full length gene.
15
5. The nucleic acid construct of claim 1, wherein the further encodes a fusion protein.
20
6. The nucleic acid construct of claim 1, wherein the RG1 polypeptide is encoded by an RG1 polynucleotide sequence.
20
7. The nucleic acid construct of claim 6, wherein the RG1 polypeptide is encoded by a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).
25
8. The nucleic acid construct of claim 1, wherein the RG2 polypeptide is encoded by an RG2 polynucleotide sequence.
30
9. The nucleic acid construct of claim 8, wherein the RG2 polypeptide is encoded by a polynucleotide sequence selected from the group consisting of: SEQ ID NO:21 (RG2A);

SEQ ID NO:23 (RG2B); SEQ ID NO:25 (RG2C); SEQ ID NO:27 (RG2D); SEQ ID
NO:29 (RG2E); SEQ ID NO:31 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:35
(RG2H); SEQ ID NO:37 (RG2I); SEQ ID NO:39 (RG2J); SEQ ID NO:41 (RG2K); SEQ
ID NO:43 (RG2L); SEQ ID NO:45 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89
5 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D);
SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID
NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107
(RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L);
SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID
10 NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126
(RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132
(RG2W).

10. The nucleic acid construct of claim 1, wherein the RG3 polypeptide is encoded by
15 an RG3 polynucleotide sequence.

11. The nucleic acid construct of claim 10, wherein the RG3 polypeptide is encoded by
a polynucleotide sequence as set forth in SEQ ID NO:68.

20 12. The nucleic acid construct of claim 1, wherein the RG4 polypeptide is encoded by
an RG4 polynucleotide sequence.

13. The nucleic acid construct of claim 12, wherein the RG4 polypeptide is encoded by
25 a polynucleotide sequence as set forth in SEQ ID NO:69.

14. The nucleic acid construct of claim 1, wherein the RG5 polypeptide is encoded by
an RG5 polynucleotide sequence.

20 15. The nucleic acid construct of claim 14, wherein the RG5 polypeptide is encoded by
a polynucleotide sequence as set forth in SEQ ID NO:134.

16. The nucleic acid construct of claim 1, wherein the RG7 polypeptide is encoded by an RG7 polynucleotide sequence.

17. The nucleic acid construct of claim 16, wherein the RG7 polypeptide is encoded by 5 a polynucleotide sequence as set forth in SEQ ID NO:136.

18. The nucleic acid construct of claim 1, further comprising a promoter operably linked to the RG polynucleotide.

10 19. The nucleic acid construct of claim 18, wherein the promoter is a plant promoter.

20. The nucleic acid construct of of claim 19, wherein the plant promoter is a disease resistance promoter.

15 21. The nucleic acid construct of claim 19, wherein the plant promoter is a lettuce promoter.

22. The nucleic acid construct of claim 18, wherein the promoter is a constitutive promoter.

20

23. The nucleic acid construct of claim 18, wherein the promoter is an inducible promoter.

25 24. The nucleic acid construct of claim 18, wherein the promoter is a tissue-specific promoter.

25. A nucleic acid construct comprising a promoter sequence from an RG gene linked to a heterologous polynucleotide.

30 26. A transgenic plant comprising a recombinant expression cassette comprising a promoter operably linked to an RG polynucleotide.

27. The transgenic plant of claim 26, wherein the plant promoter is a plant promoter.
28. The transgenic plant of claim 26, wherein the plant promoter is a viral promoter.
- 5 29. The transgenic plant of claim 26, wherein the plant promoter is a heterologous promoter.
30. The transgenic plant of claim 26, wherein the plant is lettuce.
- 10 31. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:1 (RG1A), SEQ ID NO:2 (RG1B), SEQ ID NO: 3 (RG1C), SEQ ID NO:4 (RG1D), SEQ ID NO:5 (RG1E), SEQ ID NO:6 (RG1F), SEQ ID NO:7 (RG1G), SEQ ID NO:8 (RG1H), SEQ ID NO:9 (RG1I), and SEQ ID NO:10 (RG1J).
- 15 32. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:21 (RG2A); SEQ ID NO:23 (RG2B); SEQ ID NO:25 (RG2C); SEQ ID NO:27 (RG2D); SEQ ID NO:29 (RG2E); SEQ ID NO:31 (RG2F); SEQ ID NO:33 (RG2G); SEQ ID NO:35 (RG2H); SEQ ID NO:37 (RG2I); SEQ ID NO:39 (RG2J); SEQ ID NO:41 (RG2K); SEQ ID NO:43 (RG2L); SEQ ID NO:45 (RG2M); SEQ ID NO:87 (RG2A); SEQ ID NO:89 (RG2B); SEQ ID NO:91 (RG2C); SEQ ID NO:93 (RG2D) and SEQ ID NO:94 (RG2D); SEQ ID NO:96 (RG2E); SEQ ID NO:98 (RG2F); SEQ ID NO:100 (RG2G); SEQ ID NO:102 (RG2H); SEQ ID NO:104 (RG2I); SEQ ID NO:106 (RG2J) and SEQ ID NO:107 (RG2J); SEQ ID NO:109 (RG2K) and (SEQ ID NO:110 (RG2K); SEQ ID NO:112 (RG2L); SEQ ID NO:114 (RG2M); SEQ ID NO:116 (RG2N); SEQ ID NO:118 (RG2O); SEQ ID NO:120 (RG2P); SEQ ID NO:122 (RG2Q); SEQ ID NO:124 (RG2S); SEQ ID NO:126 (RG2T); SEQ ID NO:128 (RG2U); SEQ ID NO:130 (RG2V); and, SEQ ID NO:132 (RG2W).
- 25 33. The transgenic plant of claim 26, wherein the RG polynucleotide is selected from the group consisting of SEQ ID NO:68 (RG3) and SEQ ID NO:69 (RG4).

34. The transgenic plant of claim 26, wherein the RG polynucleotide comprises a sequence as set forth in SEQ ID NO:134 (RG5).

5 35. The transgenic plant of claim 26, wherein the RG polynucleotide comprises a sequence as set forth in SEQ ID NO:136 (RG7).

10 36. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG1 polypeptide selected from the group consisting of SEQ ID NO:11 (RG1A), SEQ ID NO:12 (RG1B), SEQ ID NO: 13 (RG1C), SEQ ID NO:14 (RG1D), SEQ ID NO:15 (RG1E), SEQ ID NO:16 (RG1F), SEQ ID NO:17 (RG1G), SEQ ID NO:18 (RG1H), SEQ ID NO:19 (RG1I), and SEQ ID NO:20 (RG1J).

15 37. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG2 polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O); SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

20 38. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG3 polypeptide with a sequence as set forth by SEQ ID NO:138.

30 39. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG4 polypeptide with a sequence as set forth by SEQ ID NO:139.

40. The transgenic plant of claim 26, wherein the RG polynucleotide encodes an RG5 polypeptide with a sequence as set forth by SEQ ID NO:135.

41. A method of enhancing disease resistance in a plant, the method comprising
5 introducing into the plant a recombinant expression cassette comprising a promoter functional in the plant and operably linked to an RG polynucleotide sequence.

42. The method of claim 41, wherein the plant is a lettuce plant.

10 43. The method of claim 41, wherein the RG polynucleotide encodes an RG polypeptide selected from the group consisting of SEQ ID NO:22 and SEQ ID NO:41 (RG2A); SEQ ID NO:24 and SEQ ID NO:42 (RG2B); SEQ ID NO:43 (RG2C); SEQ ID NO:44 (RG2D); SEQ ID NO:45 (RG2E); SEQ ID NO:46 (RG2F); SEQ ID NO:47 (RG2G); SEQ ID NO:48 (RG2H); SEQ ID NO:49 (RG2I); SEQ ID NO:50 (RG2J); SEQ ID NO:51 (RG2K); SEQ ID NO:52 (RG2L); SEQ ID NO:53 (RG2M); SEQ ID NO:88 (RG2A); SEQ ID NO:90 (RG2B); SEQ ID NO:92 (RG2C); SEQ ID NO:95 (RG2D); SEQ ID NO:97 (RG2E); SEQ ID NO:99 (RG2F); SEQ ID NO:101 (RG2G); SEQ ID NO:103 (RG2H); SEQ ID NO:105 (RG2I); SEQ ID NO:108 (RG2J); SEQ ID NO:111 (RG2K); SEQ ID NO:113 (RG2L); SEQ ID NO:115 (RG2M); SEQ ID NO:117 (RG2N); SEQ ID NO:119 (RG2O);
15 20 SEQ ID NO:121 (RG2P); SEQ ID NO:123 (RG2Q); SEQ ID NO:125 (RG2S); SEQ ID NO:127 (RG2T); SEQ ID NO:129 (RG2U); SEQ ID NO:131 (RG2V); and, SEQ ID NO:133 (RG2W).

44. The method of claim 41, wherein the RG polynucleotide encodes an RG polypeptide selected from the group consisting of SEQ ID NO:138 (RG3); SEQ ID NO:139 (RG4); and
25 SEQ ID NO:135 (RG5).

45. The method of claim 41, wherein the promoter is a tissue-specific promoter or a plant disease resistance promoter.

46. The method of claim 41, wherein the promoter is a constitutive promoter or an inducible promoter.

47. A method of detecting RG resistance genes in a nucleic acid sample, the method comprising:

contacting the nucleic acid sample with an RG polynucleotide to form a hybridization complex; and,

wherein the formation of the hybridization complex is used to detect the RG resistance gene in the nucleic acid sample.

10

48. The method of claim 47, wherein the RG polynucleotide is an RG1 polynucleotide.

49. The method of claim 47, wherein the RG polynucleotide is an RG2 polynucleotide.

15

50. The method of claim 47, wherein the RG polynucleotide is an RG3 polynucleotide, an RG4 polynucleotide, an RG5 polynucleotide or an RG7 polynucleotide.

51. The method of claim 47, wherein the RG resistance gene is amplified prior to the step of contacting the nucleic acid sample with the RG polynucleotide.

20

52. The method of claim 51, where the RG resistance gene is amplified by the polymerase chain reaction.

53. The method of claim 47, wherein the RG polynucleotide is labeled.

25

54. An RG polypeptide having at least 60% sequence identity to a polypeptide selected from the group consisting of: an RG1 polypeptide, an RG2 polypeptide, an RG3 polypeptide, an RG4 polypeptide, an RG5 polypeptide, and an RG7 polypeptide.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US98/00615

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.

US CL :435/6, 91.2, 418, 419; 530/350; 536/23.1, 23.6, 24.1; 800/205

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.2, 418, 419; 530/350; 536/23.1, 23.6, 24.1; 800/205

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PARAN et al. Development of Reliable PCR-Based Markers Linked to Downy Mildew Resistance Genes in Lettuce. Theor. Appl. Genet. 1993. Vol. 85, No. 8, pages 985-993, see entire article.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	KESSELI et al. Analysis of a Detailed Genetic Linkage Map of Lactuca sativa (Lettuce) Constructed From RFLP and RAPD Markers. Genetics. April 1994. Vol. 136, No. 4, pages 1435-1446, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	MICHELMORE, RW. Isolation of Disease Resistance Genes from Crop Plants. Current Opinion in Biotechnology. 1995. Vol. 6, No. 2, pages 145-152, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54

 Further documents are listed in the continuation of Box C. See patent family annex.

Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
B earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"A"	document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means		
P document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
14 MARCH 1998

Date of mailing of the international search report

13 APR 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230Authorized officer
PHUONG BUI
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00615

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PARAN et al. Recent Amplification of Triose Phosphate Isomerase Related Sequences in Lettuce. Genome. 1992. Vol. 35, No. 4, pages 627-635, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54
Y	PARAN et al. Identification of Restriction Fragment Length Polymorphism and Random Amplified Polymorphic DNA markers linked to Downy Mildew Resistance Genes in Lettuce, Using Near-Isogenic Lines. Genome. 1991. Vol. 34, No. 6, pages 1021-1027, see entire document.	1-6, 8, 10, 12, 14, 16, 18-30, 41-42, 45-54

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US98/00615

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.: 7, 9, 11, 13, 15, 17, 31-40, 43-44 because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

these claims are drawn to numerous sequences identified by SEQ ID NOs. However, since no computer readable form was submitted, no meaningful search could be carried out.

3. Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/00615

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (6):

A01H 1/00; C07H 21/04; C07K 14/00; C12N 5/04, 5/10; C12P 19/34; C12Q 1/68